Roles for Terminal Uridyl Transferases in the Post-Transcriptional Regulation of Developmental miRNAs

Roles for Terminal Uridyl Transferases in the Post-Transcriptional Regulation of Developmental miRNAs

A dissertation presented

by

James Edward Thornton

to

The Committee on Higher Degrees in Biological Sciences in Public Health

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Biological Sciences in Public Health

Harvard University

Cambridge, Massachusetts

September, 2013

Dissertation Advisor: Dr. Richard Gregory James Edward Thornton

Roles for Terminal Uridyl Transferases in the Post-Transcriptional Regulation of Developmental miRNAs

Abstract

MicroRNAs (miRNAs) are a diverse and evolutionarily conserved class of non-coding RNAs that play a multitude of roles in many branches of eukaryotic biology. The regulation of miRNAs is dynamically controlled both spatially and temporally, and the expression of miRNAs can be modulated at the level of transcription or at points downstream of the miRNA maturation process. A relevant example of post-transcriptional miRNA regulation is the blockade of let-7 precursor miRNAs by Lin28 in embryonic stem cells. This pathway, which is initiated by the small RNA-binding protein Lin28, recruits the terminal uridyl transferase (TUTase) Zcchc11 to add a non-templated oligouridine tail to the miRNAs 3' end, and signals it for degradation by the cytoplasmic exonuclease Dis3l2. The Lin28/let-7 axis is essential for development and metabolic homeostasis, and is reactivated in a subset of human cancers. This thesis describes the biochemical mechanism underlying Lin28-mediated degradation of let-7, as well as a novel role for Zcchc11 and the related TUTase Zcchc6 in targeting mature developmental miRNAs in a Lin28-independent manner.

As shown in Chapter 2, we uncovered the mechanism through which the multi-domain protein Zcchc11 recognizes pre-let-7 and Lin28. This work used *in vitro* biochemical techniques to monitor the activity of Zcchc11 mutants and uncovered residues that are required to form the

ternary complex of Zcch11, Lin28, and let-7. This work also found the highly similar TUTase Zcchc6 to function redundantly with Zcchc11 *in vitro* and *in vivo.*

Chapter 3 describes the discovery of an intrinsic sequence preference for Zcchc6 and Zcchc11 towards single stranded mature miRNAs. This sequence motif, which is necessary and sufficient to mediate TUTase-miRNA binding and monouridylation, was surprisingly found only in developmental miRNAs that preferentially target Hox genes. Bioinformatic approaches and knockdown experiments in cultured cells showed that indeed these TUTases preferentially target the motif-containing miRNAs *in vivo*. Surprisingly, after observing a specific TUTase-mediated loss of uridylation, these same miRNAs underwent a proportional increase of non-templated uridylation. Furthermore, both Zcchc6 and Zcch11 were found to be absent from nearly all adult tissues and were potently down-regulated during an *in vitro* model of differentiation, suggesting that their activity is primarily relevant in early stages of development.

Considered together, these data refine the mechanism of known TUTase activity and also expand the roles for TUTases in development. This work should be useful in the future to assist in the development of small molecule inhibitors against this relevant oncogenic pathway, and to elucidate the myriad levels of gene regulation during development.

ACKNOWLEDGEMENTS

None of the work described here could have been accomplished without a tremendous team of supporters and collaborators. First and foremost I must thank Richard Gregory for his tireless support and mentorship, from the beginning of my rotation to the final days of my Ph.D. I also must thank everyone else in the lab for their guidance and insight, from the most mundane issues to the highly complicated. I especially want to note those who predated me in the Gregory lab and who were constantly present during a large part of my studies; Elena Piskounova, Hao Ming Chang, Robinson Triboulet, and Robert Lapierre. For those who came after me my debt is nearly as large; Natalia Martinez, Masaki Mori, Peng Du, and Shuibin Lin. I also had the amazing opportunity to mentor a number of summer students and rotation students, and this experience affirmed to me the value of teaching and its central role in my own education, so special thanks to Adam Wilkinson, Femke Feringa, SuFei Yue, Elena Grossi, and Jeff Davis.

Over these years our lab has depended on several other research groups, without whom much of our work would not be possible. Thanks especially to the laboratories of George Daley (and the Lin28 subgroup of Hao Zhu, Shyh-Chang Ng, John Powers, Kaloyan Tsanov, Daisy Robinton, and Michael Kharas), Lenoard Zon (especially Lili Jing), and Dimitrios Iliopoulos. I must also mention the important guidance—however intense—of my Dissertation Advisory Committee (DAC) composed of Marianne Wessling-Resnick (chair), Kevin Struhl, Steve Buratowski, and Tiffany Horng. The administrative staffs at both Children's Hospital Boston and the Harvard School of Public Health offered tremendous non-scientific support, and I would be remiss to not mention their support as well.

Five years is a long time and despite my best efforts it was not all spent at the bench. There was a lot of time taken to decompress, discuss, and even have fun and I can't overstate the

central role that my friends and family played in this process. Academically, the friends I made at Science in the News and the Science Policy PATH program were a perfect supplement to a scientific life, and the many other graduate students I have gotten to know on various levels have been invaluable in maintaining a balanced social life. The role of my mother and father in supporting me and encouraging me to reach this point is truly the one common thread that has been unwavering, yet I can't forget the first time I was introduced to the scientific method by my late grandfather over two decades ago. Finally, I must mention the daily support offered by Nadza Durakovic, and the importance she has shown me of keeping things in perspective during the turbulent times of graduate life.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

Introduction and Background

Text and figures partially adapted from "How does Lin28 let-7 control development and disease?" by J.E. Thornton and R.I. Gregory (2012). The author created all figures, except where indicated otherwise.

Thornton and Gregory. How does Lin28 let-7 control development and disease?. Trends Cell Biol (2012) vol. 22 (9) pp. 474-82

Introduction

As the field of molecular biology blossomed in the middle part of the 20th century, the first mechanistic insight was gained on the inner-workings of the cell. New forms of experimentation, often co-opting principles from the physical sciences, led to completely novel paradigms that still form the bases of biology today. The rapid expansion of this new field, as fundamental principles were developed, formed the foundations of molecular biology and allowed for the first time a deeper understanding of the mechanistic nature of life.

As the field progressed into the future, many of these findings were supported by further experimentation, but many important revisions have since been made that complicate the basic rules first outlined decades ago. The simplicity of the operon as described by Jacob and Monod (1961), while elegant, is not a general feature of all cellular life. The importance of DNA-based inheritance certainly explains a great deal of human traits and diseases, but the importance of epigenetics—non-sequence-based regulators of gene expression and inheritance—are increasingly difficult to overstate. The human genome is copied in a semi-conservative manner and RNA synthesis is largely based on a DNA template as described by Watson and Crick (1953), but non-templated nucleotide addition is widespread in animal genomes, and RNA-dependent RNA synthesis is an essential activity in many processes, both cellular and viral. Finally, the so-called Central Dogma of Molecular Biology, which states that DNA becomes RNA, which then becomes protein, has been heavily revised since its proposal in 1970 (Crick, 1970). A major implication of the Central Dogma is that the ultimate role of a gene is to become a protein and elicit a function. Because of work in the last several decades we now know that a significant proportion of the genome is transcribed, with much of it remaining as stable non-coding RNAs (Reviewed in Jacquier, 2009). While many of these transcripts appear to have no

functional role in gene expression or remain elusive in their activity, there are several subclasses of non-coding RNAs that are central to maintaining cellular homeostasis and regulating dynamic processes including development and oncogensis.

The idea that RNA may play a regulatory role independent of protein synthesis is not new. As early as 1969 Britten and Davidson theorized that external signals could trigger the synthesis of RNA molecules, which would then be sensed by internal receptors capable of feeding forward to later rounds of mRNA production (Britten & Davidson, 1969). Over the next several decades, long RNAs such as H19 and Xist were identified and appeared to lack any large open reading frames (ORFs) and could be retained in the nucleus precluding the possibility of translation (Brockdorff et al., 1992, Brannan et al., 1990). Shortly thereafter small RNAs of less than 25 nucleotides (nt) were identified as developmental regulators in the nematode *C. elegans* and as silencers of viral genes in plants (Lee et al., 1993; Wightman et al., 1993; Ratcliff et al., 1997; Hamilton & Baulcombe, 1999). Work from the labs of Victor Ambros and Gary Ruvkun showed that the developmental timing, or "heterochronic" gene lin-4 did not contain any large open reading frames (ORFs) and that rather than existing as a stable mRNA, was processed to a discrete 22nt product. Intriguingly, a portion of this RNA was complementary to sites in the 3'UTR of another heterochronic gene, lin-41, and this pairing was necessary and sufficient to direct temporal silencing of lin-41 (Lee et al., 1993; Wightman et al., 1993). These short regulatory RNAs were expressed in a manner that was inversely proportional to the levels of mRNAs carrying complementary sites, suggesting they could be effector molecules responsible for gene silencing. These discoveries in plants and animals laid the groundwork for the general concept of post-transcriptional gene silencing (PTGS), a process initially thought to play a minor role in normal cellular processes.

A breakthrough in elucidating the mechanism of PTGS came in 1998 when Andrew Fire and colleagues discovered that long double-stranded RNA could potently silence genes that shared extensive homology when they were injected into *C. elegans* embryos (Fire et al. 1998). Importantly, this silencing was efficient even when the RNA molecules existed at only a few copies per cell, implying the presence of a multiple turnover complex driving gene silencing, rather than a simple stoichiometric interruption of mRNA function. The implications of this discovery were unclear and pointed towards a mechanism that could regulate exogenous RNAs, rather than one that could play a role in regulating gene expression through endogenous small RNAs.

Several years later a second 22nt developmental RNA was identified in *C. elegans*. This molecule, let-7, shared many characteristics with lin-4 including its temporal expression as well as its targeting and silencing of other developmental genes through their 3'UTRs (Reinhart et al., 2000). let-7 targeted several mRNAs not previously known to be regulated post-transcriptionally, yet it shared target mRNAs with the lin-4 small RNA as well, including lin-41 and lin-28. Surprisingly, later that year the same group found that the let-7 sequence and expression pattern were conserved across animal phyla from mollusks to humans (Pasquinelli et al., 2000). In rapid successions at least three independent groups discovered a broad class of these small regulatory RNAs that vary in their sequence, expression pattern, and conservation. (Lagos-Quintana et al., 2001, Lau et al., 2001, Lee & Ambros, 2001). Termed microRNAs (miRNAs), this gene class was annotated and organized into families based on predicted target mRNAs. Although some miRNAs may be somewhat similar in sequence to other miRNA genes, the primary criteria for classifying miRNAs was quickly understood to be the so-called seed sequence: nucleotides 2-8

of mature miRNA molecules, which is the primary determinant of target specificity and can predict most of the repressive action of a given miRNA (Lewis et al., 2003).

The Machinery of miRNA Biogenesis

Although both small interfering RNAs (siRNAs) and miRNAs were known to repress complementary genes, the biochemical pathways underlying their activities was initially unclear. While siRNAs can be processed from long exogenous dsRNAs or introduced as short duplexes, primary miRNA transcripts—so-called pri-miRNAs—are initially transcribed by RNA polymerase II. They have one or more hairpins containing mature miRNAs and long ssRNA flanking sequences up and downstream of the central miRNA hairpin. They contain a 5' methylguanine cap and 3' poly(A) tail and are processed in the nucleus to yield a miRNA precursor or pre-miRNA, before further processing to the characteristic 22nt duplex (Lee et al., 2002). The initial processing event in miRNA synthesis is through the nuclear Microprocessor complex. The minimal Microprocessor is composed of the enzyme Drosha, as well as its accessory factor Pasha in *D. melanogaster* and *C. elegans* or DGCR8 (DiGeorge Syndrome critical region 8) in humans and other organisms (Gregory et al., 2004; Denli et al., 2004). Drosha was identified through a biochemical candidate approach and belongs to the family of RNase-III enzymes (Lee et al., 2003), a group of enzymes present in both prokaryotes and eukaryotes that are known to cleave dsRNA. RNase-III family members are characterized by their ability to produce precise cleavage products with 2nt 3' overhangs, which contain 5' phosphates and 3' hydroxyl groups.

Much of the specificity in pri-miRNA cleavage is determined by the miRNA itself. Lee et al. (2002) showed that structural modifications affecting pri-miRNA folding perturbed the

double stranded nature of the miRNA stem and could abolish pre-miRNA production, indicating that a dsRNA pri-miRNA stem is a preferred Drosha substrate. Several groups also showed that the terminal loop is dispensable for pri-miRNA cleavage and that the single stranded flanking regions play a significant role in proper processing, likely due to DGCR8 binding this ssRNA/dsRNA junction. (Lee et al., 2002; Lee et al., 2003; Zeng & Cullen, 2005; Zeng et al., 2005). The length of the pri-miRNA stem is also important in Microprocessor recognition, and its structure directly contributes to pri-miRNA cleavage products. DGCR8 binds to the junction between the single and double stranded regions of the pri-miRNA stem loop and induces cleavage ~11nt, or one helical turn from this junction (Lee et al., 2003). Together, the two essential components of the Microprocessor form a precise complex capable of recognizing and processing miRNAs in a sequence-nonspecific manner.

Finally, it is important to note that many accessory proteins have been identified as associating with the Microprocessor in large-scale immunoprecipitations (Chendrimada et al., 2005). Many of these factors have not been found to play a role in miRNA biogenesis and may indicate alternative cellular roles for Drosha or a further level of complexity in miRNA processing that has yet to be fully understood.

Since exogenous complementary dsRNAs were initially shown to elicit gene silencing in a manner somewhat similar to miRNAs, some assumed that the two pathways shared common processing pathways (Tuschl et al., 2000; Zamore et al., 2000; Hutvagner et al., 2001; Hutvagner & Zamore, 2002). The prime characteristic shared between these two types of RNA species was, after all, a precise series of nucleolytic cleavages that yielded ~22nt dsRNAs. This catalytic activity was eventually attributed to the enzyme Dicer, which is highly conserved across evolutionary taxa. Like Drosha, Dicer is a member of the RNase-III family. A candidate

approach using purified Dicer showed that it had specific, yet sequence-independent activity capable of producing ~22nt dsRNAs and that this activity resided in a high molecular weight complex (Bernstein et al., 2001; Hammond et al., 2000). These results indicated that the cleavage of long RNAs to yield siRNAs was a distinct activity that likely occurred upstream of target RNA silencing. Dicer was also shown to mediate the cleavage of miRNAs, thus providing the first evidence of overlap between the RNAi and miRNA pathways, initially believed to be distinct (Hutvagner et al., 2001).

Dicer is the central member of Class III RNase-III enzymes. It is characterized by its tandem RNase-III domains and dsRNA-binding domain, all of which are positioned at the C-terminus of the protein. Dicer also contains an N-terminal helicase domain of the DExDc variety that is involved in enhancing the processing of thermodynamically unstable RNA substrates (Soiffer et al., 2008; Kim et al., 2008). Another crucial domain for catalytic function is the PAZ (Piwi/Argonaute/Zwille) domain, which binds 3' OH groups of ssRNA (Bernstein et al., 2001; Hutvagner et al., 2001). Together, these domains provide the necessary features for dsRNA recognition and cleavage.

The catalytic core of Dicer is composed of its two RNase-III domains. In prokaryotes, RNase-III-containing enzymes form homodimers, which are capable of cleaving a broad spectrum of dsRNAs. The two RNase-III domains in Dicer were initially believed to act independently, with one domain cleaving both strands of the RNA duplex at one point, while the other cleaved the same duplex 22nt away. It was later shown that instead, both bacterial RNase-III homodimers and Dicer contained only a single catalytic center that was capable of cleaving dsRNA at a single point in an ATP-independent manner. Specifically, the N-terminal-most RNase-III cleaves the strand that will contain the free 5' end, while the C-terminal RNase-III

produces the free 3' end (Zhang et al., 2002; Zhang et al., 2004; MacRae et al., 2006). These paired domains in Dicer are offset such that a dsRNA with a 2nt 3' overhang is produced. (Zhang et al., 2002; Zhang et al., 2004). If Dicer contains only a single processing center, how does it faithfully cleave RNA duplexes in 22nt intervals? This question was addressed through structural studies of *Giardia intestinalis* Dicer. The PAZ domain of this Dicer family member is positioned distal to the intramolecular RNase-III dimer and the distance from the PAZ domain to the catalytic center is roughly 22nt, or the length of two helical turns of an RNA duplex. The region between the PAZ domain and the paired RNase-III domains is lined with basic residues, which create an ideal surface for binding nucleic acids. Furthermore, the PAZ domain, which binds free 3' hydroxyl groups, (Ma et al., 2004) was shown to recognize the staggered ends of dsRNAs previously processed by Dicer or other RNase-III enzymes such as the miRNA-specific enzyme Drosha. This structural arrangement implies that Dicer enzymes can only efficiently act on substrates with free termini and must process long dsRNA in a step-wise fashion. These findings show that Dicer acts as a molecular ruler, positioning its dsRNA substrates so that they can be accurately cleaved to produce ~22nt products.

Although the general rules of Dicer processing were gleaned from bacterial and lower eukaryotic enzymes, a recent study on human Dicer revealed a novel structure-function relationship that is essential to miRNA biogenesis. Examining processing products from RNA duplexes with long 3' overhangs, Park et al. (2011) discovered that Dicer still primarily cleaved these substrates to 22nt products, implying that 3' counting is not sufficient to explain Dicer activity. Indeed the authors were able to identify a basic patch that recognizes 5' phosphate moieties of RNA, and in conjunction with PAZ-mediated 3' hydroxyl recognition, cleaves the RNA duplexes ~22nt from the 5' end. Interestingly, this preference is conserved in *Drosophila*

Dicer, but is absent in *Giardia* Dicer. The latter enzyme had been studied extensively and a solution structure was solved, misleading several groups to conclude that 5' processing was generally unimportant in the Dicer processing pathway. This study nicely uncovered the complexity and diversity that forms some of the most basic steps in miRNA processing pathways across divergent evolutionary groups.

The functional center of both miRNA- and siRNA-mediated gene repression is the RNA-Induced Silencing Complex (RISC). In this multiprotein complex, catalysis is carried out by members of the Argonaute (Ago) protein family. The initial identification of genes encoding Ago family members was performed in *Arabidopsis* when a screen uncovered several loci responsible for proper leaf morphology (Bohmert et al., 1997). Soon thereafter, Ago homologs were identified through *C. elegans* mutant screens, which showed them to be essential for RNAi (Tabara et al., 1999). Biochemical studies that attempted to purify RISC activity using *Drosophila* embryos, or phylogenetic studies that compared effector molecules from disparate organisms identified Ago proteins as a central component of functional RISCs (Catalanotto et al., 2000; Hammond et al., 2001). Ago proteins vary widely in copy number depending on organism. The fission yeast *S. pombe* encodes only a single Ago and the human genome contains at least five Ago family members, while the nematode *C. elegans* contains more than twenty-five distinct Ago genes. The roles of these seemingly redundant proteins are not fully understood, and it is unclear if all share equal importance in processing endogenous small RNAs.

Argonaute proteins are highly conserved and characterized by stereotypical domain architecture. Similar to Dicer, Ago proteins contain a PAZ domain involved in binding 3' ends of single stranded RNA. Determined by the crystal structure of an archaeal Argonaute, the PAZ domain is located away from the rest of the protein on a stalk, ideally suited for binding the

termini of small RNAs. The amino acid residues preferentially located for nucleic acid binding are almost identical in position and number between archaea and mammals. Later studies, however, showed the PAZ domain to be dispensable for Ago catalysis and instead, the Mid domain, which binds to the 5' end of incorporated small RNAs, is critical for proper RISC activity (Wang et al., 2008; Wang et al., 2008a). These structural studies also led to proposals for the mechanism of RISC catalysis. The PIWI domain is unique to Ago proteins and is made up of bundled beta sheets forming a groove that was modeled to bind the length of an RNA molecule. This domain forms a classic RNase H fold with a conserved catalytic triad. Although RNase H typically cleaves RNA/DNA duplexes, the presence of this motif in a RISC-associated protein made Ago an obvious candidate for the central enzyme in RISC activity. Mutational analyses and further structural studies confirmed the catalytic role of the RNase H domain, indicating the importance of Argonaute proteins in small RNA-mediated silencing (Liu et al., 2004). To determine which, if any, of the Ago proteins was responsible for cleaving siRNA-targeted mRNAs—so-called Slicer activity—biochemical studies were performed to assay individual Ago family members for this catalytic activity *in vitro*. Mammalian Ago2 was the only species found to facilitate the cleavage of RNA duplexes, even though the other Ago proteins associated with the same miRNAs and siRNAs at similar levels (Liu et al., 2004). This observation defined the centrality of these conserved proteins in numerous realms of biology. Indeed, mice deficient for Ago2 exhibited embryonic lethality while cells expressing a catalytically inactivated Ago2 failed to mount an RNAi response (Cheloufi et al., 2010; Cifuentes et al., 2010).

The activity of RISC is required to be very precise in that it must faithfully process only RNA substrates involved in gene silencing. The small RNAs incorporated into RISC are characterized by their termini containing 2nt 3' overhangs and it was not initially clear how—or

whether—RISC selectively incorporated only the proper guide strand to target cognate mRNAs. The fidelity of RISC is very high and a mechanism either intrinsic to the enzyme or to the RNAs was believed to exist. Since the guide strand, or the strand of RNA directly involved in silencing, can belong to either half of the symmetrical duplex, a selection mechanism mediated by RISC itself was not obvious. Thermodynamic studies soon revealed that strand bias is largely determined by the relative stability of the two ends of the RNA duplex. A computational study showed that the guide strand of annotated miRNAs preferentially contained a 5' end with weaker thermodynamic stability, which perhaps facilitated their separation from the "passenger" strand and subsequent incorporation into RISC. Indeed, siRNAs could be engineered with varying 5' thermodynamic stabilities and shown to elicit a potent RNAi response depending on which strand was selected by RISC. This detection mechanism is sensitive enough to interpret relative thermodynamic differences as small as a single hydrogen bond (Khvorova et al., 2003; Schwarz et al., 2003). Therefore, both miRNAs and siRNAs are asymmetrically processed to incorporate a dominant ssRNA species. This strong bias found in most miRNAs indicates that the 5' ends of mature miRNAs are under strong selective pressure, which could either be responsible for or due to the requirements for conserving the miRNA seed sequence.

Two recent crystal structures of human Argonaute 2 in complex with small RNAs support many of the initial findings on Argonaute activity (Schirle et al., 2012; Elkayam et al., 2012). There is remarkable homology between archaeal and mammalian Ago proteins despite billions of years of evolutionary divergence, down to the same catalytic residues and general domain organization. Guide strands loaded into Ago2 contact nearly every functional domain, and act to stabilize the lobed protein structure, almost as a molecular "staple". Interestingly, recent independent studies have found there is a delicate balance between miRNA and Ago2

levels, where the protein is destabilized in the absence of mature miRNAs, and miRNAs are reduced after Ago2 depletion (Gibbings et al., 2012; Martinez et al., 2013). This co-relationship along with the significant conservation over evolutionary times underscores the intimate relationship between the drivers of PTGS and their substrates, as well as the importance of Argonaute proteins, even if their role in lower organisms is largely unknown.

Argonaute proteins are only a single component of a functional RISC and they are sufficient for slicer activity (Rivas et al., 2005). However, other protein components in RISC play essential roles. The first identified Ago-associated RISC component was R2D2 in flies (Liu et al., 2003). Biochemical purifications aimed at purifying Dicer-associated factors identified the small RNA-binding protein R2D2 associating tightly with Dicer activity. R2D2 was shown to be dispensable for the production of siRNAs by Dicer, but essential for RNAi. Furthermore, R2D2 was only shown to be associated with Dicer-2, the Drosophila enzyme that mediates RNAi but not miRNA-mediated gene repression. These characteristics of R2D2 led to its implication in mediating RISC loading. Later studies showed R2D2 to preferentially bind the end of siRNA duplexes that were *more* double stranded in nature, thereby facilitating the exclusion of the passenger strand while enhancing faithful RISC loading (Tomari et al., 2004). A human protein with similar characteristics to R2D2 was later identified through large-scale purifications (Chendrimada et al., 2005; Gregory et al., 2005). This protein, TRBP, shared no obvious sequence homology with R2D2 but like the fly protein physically coupled Dicer to RISC, was not essential for Dicer-mediated siRNA/miRNA generation, yet was necessary for RISC activity. Moreover, reconstitution of RISC loading and activity was accomplished by Dicer, TRBP, and Ago2 (Gregory et al., 2005; MacRae et al., 2008). This complex was capable of multiple mRNA cleavage events and functioned independent of ATP. The proteins R2D2 and TRBP are integral

components of RISC and enhance the function of small RNA-mediated gene repression. These important accessory factors significantly enhance activity through coupling distinct processing events and therefore further underscore the complexity of RISC *in vivo*.

The discrete compartmentalization of miRNA processing factors ensures proper spatial and temporal mechanistic control of miRNA biogenesis. For each step of the pathway, modular components leave characteristic marks on the RNA substrates, leaving them optimized for downstream roles. This cascade of activity has evolved over millions of years as a precise and diverse modulator of gene expression. As a result it is perhaps unsurprising that perturbations in miRNA processing components can exacerbate sensitized cell lines and promote tumorogenesis and metastasis in cultured cells and animal models (Kumar et al., 2007; Kumar et al., 2009). Indeed, a global reduction of miRNA expression is observed in many cancers and undifferentiated cells lacking Dicer or DGCR8 fail to express mature miRNAs and have differentiation defects (Kanellopoulou et al., 2005; Lu et al., 2005; Murchison et al., 2005; Wang et al., 2007). This delicate balance of regulating miRNAs throughout an organism's developmental history is therefore central at many levels of biology.

Lin28 and let-7: A crucial developmental axis

Identified as the first mammalian miRNAs to have functional and sequence conservation to a *C. elegans* counterpart, the mammalian let-7 family comprises twelve members expressed from eight distinct loci (let-7a-1, -2, -3; let-7b; let-7c; let-7d; let-7e; let-7f-1, -2; let-7g; let-7i; miR-98) (Pasquinelli et al., 2000; Johnson et al., 2005). Although distributed throughout the genome, many members of the let-7 family are coordinately regulated during development with let-7 expressed at high levels in differentiated cell-types. During differentiation of mouse

embryonic stem cells (ESCs), mature let-7 miRNAs accumulate yet the corresponding primary let-7 transcript (pri-let-7) remains constant (Thomson et al., 2006; Suh et al., 2004). The discrepancy between let-7 transcriptional activity and the delay in mature let-7 accumulation implies the existence of a negative regulatory factor inhibiting let-7 processing in undifferentiated cells. To determine proteins that potentially mediate this block, several groups performed biochemical purifications of complexes associating with pre-let-7. Mass Spectrometric analyses revealed two predominant interacting proteins to be Lin28 and Lin28B. Lin28 family members are small (<30kDa) proteins containing two CCHC-type zinc fingers and a Cold-Shock Domain (CSD), both implicated in RNA binding, and are specifically expressed in embryonic cells (Figure 1.1B). Purified Lin28A blocks the processing of let-7 at both the pri- and pre-miRNA steps, as Lin28-bound let-7 is resistant to cleavage *in vitro* using either purified Microprocessor or Dicer complexes. Furthermore, overexpression of Lin28A or Lin28B in differentiated cells potently and specifically represses mature let-7 levels, while RNAi-mediated knockdown of Lin28A in undifferentiated cells or Lin28B in cancer cells is sufficient to specifically alleviate let-7 repression (Heo et al., 2008; Newman et al., 2008; Viswanathan et al., 2008; Rybak et al., 2008). Thus, regulation of let-7 expression during embryonic development, stem cell differentiation, and various cancers is controlled primarily through the post-transcriptional blockade of let-7 biogenesis by Lin28 proteins (Figure 1.1A).

Since Lin28A and Lin28B mRNAs are themselves let-7 targets this Lin28/let-7 axis establishes a double-negative feedback loop whereby either let-7 or Lin28 is expressed at high levels, promoting either a differentiated or embryonic cell fate, respectively (Rybak et al., 2008). Aside from their role as developmental regulators, Lin28 proteins are oncogenes reactivated in ~15% of all cancers analyzed, functioning largely through their repression of let-7 miRNAs

(Viswanathan et al., 2009). In accordance with reestablishing a gene expression signature reminiscent of ESCs, Lin28 overexpression in conjunction with three other defined factors is sufficient to reprogram somatic cells to an induced pluripotent state (Yu et al., 2007). Antagonizing let-7 function with antisense oligonucleotides similarly enhances the reprogramming of fibroblasts to induced pluripotent stem cells (iPSCs) (Takahashi & Yamanaka, 2006; Stadtfeld et al., 2010; Melton et al., 2010). Ranging from embryonic development, to cancer, to inflammation and metabolic processes, the Lin28/let-7 axis is now recognized as central to maintaining proper cell fate and coordinating proliferation, growth, and energy utilization at the cellular level as well as growth, developmental timing, tissue homeostasis, and metabolism in whole organisms. This primal pathway contributes to human disease when dysregulated as exemplified by the oncogenic role of Lin28 in a wide variety of cancers.

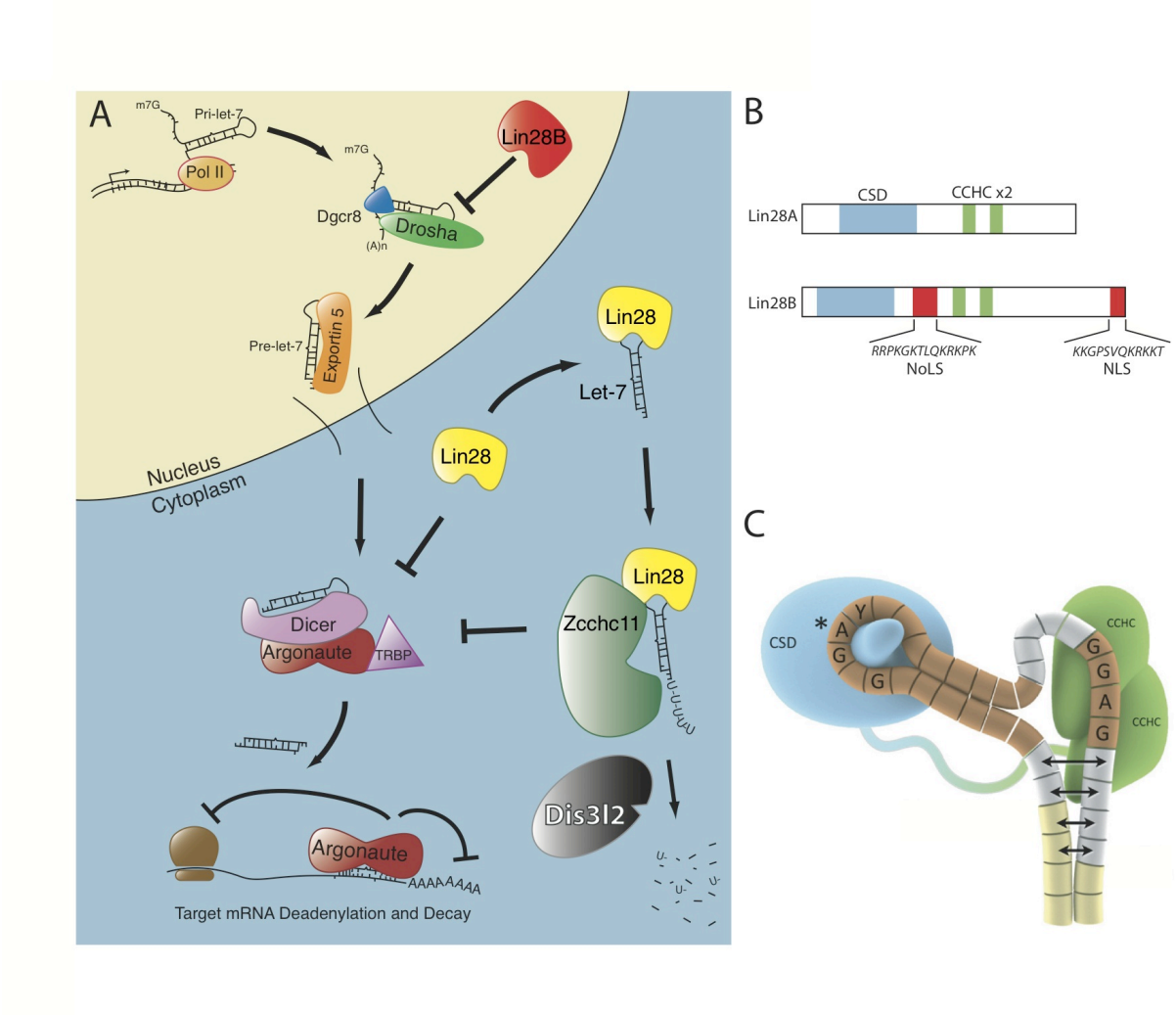


Figure 1.1: Lin28 selectively blocks let-7 biogenesis

Figure 1.1 continued

A. Model of Lin28-mediated control of let-7 biogenesis. Left: In the absence of Lin28 family members, let-7 microRNAs (miRNAs) are processed through the canonical miRNA biogenesis pathway. Transcribed by RNA polymerase II, pri-let-7 is cleaved by the microprocessor complex of DGCR8 and Drosha, yielding pre-let-7, which is exported to the cytoplasm. Pre-let-7 is then cleaved by Dicer, resulting in a mature let-7 duplex that is asymmetrically loaded into one of several Argonaute proteins within the RNA-induced Silencing Complex (RISC). RISC-mediated repression of target messenger RNAs (mRNAs) can occur through translational inhibition and deadenylation or, if the miRNA shares perfect sequence complementarity with its target, mRNA cleavage. Right: Expression of Lin28A in the cytoplasm blocks let-7 processing by Dicer through direct RNA–protein interactions with the terminal loop of pre-let-7 and recruits the terminal uridyl transferase (TUTase) Zcchc11/TUT4 to catalyze the addition of a 3' oligouridine tail, marking pre-let-7 for degradation by Dis3l2. Lin28B is expressed primarily in the nucleolus and binds to pri-let-7, thereby blocking the activity of the microprocessor complex through an unknown TUTase-independent mechanism.

B. Domains of Lin28A and Lin28B proteins. Lin28A and Lin28B share significant protein sequence identity and comprise several common domains including a Cold Shock Domain (CSD, blue) and CCHC zinc fingers (green). Lin28B contains both a nuclear localization signal (NLS) and a nucleolar localization signal (NoLS), explaining its nucleolar localization.

C. Molecular of for Lin28 interaction with pre-let-7. A cartoon representation of Lin28A binding to the terminal loop of a generic pre-let-7 miRNA based on high-resolution crystal structures. The CSD (blue) is inserted into the terminal loop, whereas the CCHC zinc fingers (green) dimerize around a conserved sequence motif (GGAG) proximal to the Dicer cleavage site. Binding of the zinc fingers partially unwinds the RNA duplex of pre-let-7, perhaps explaining the resistance of Lin28-bound pre-let-7 to Dicer cleavage. Figure reproduced from Nam et al., 2011.

Lin28 selectively inhibits let-7 biogenesis

The initial identification of Lin28 and its inverse expression with let-7 was carried out in *C. elegans*, but the molecular mechanism tying these two heterochronic genes remained unknown until quite recently. The Lin28 pathway predominantly regulates expression of the let-7 family whereas expression of other miRNAs is largely unaffected by Lin28. Lin28 binds to the terminal loop or pre-element (preE) of pre- and pri-let-7 *in vitro* and fails to repress the expression of chimeric let-7 miRNAs bearing the preE of divergent miRNAs (Piskounova et al., 2008). Indeed, the preE of let-7 alone is sufficient for Lin28 binding. These *in vitro* assays correspond well with the global effects on miRNA expression where Lin28A depletion leads to the specific accumulation of multiple let-7 miRNAs in ESCs, and similarly, Lin28A or Lin28B overexpression leads to the selective repression of let-7 miRNAs with the levels of other miRNAs remaining unchanged (Newman et al., 2008; Hagan et al., 2009; Heo et al., 2009). This binding and repression of let-7 requires both the CSD and CCHC zinc fingers of Lin28 protein (Piskounova et al., 2008; Heo et al., 2009). A more refined picture of the Lin28-let-7 interaction is provided by high-resolution crystal structures of a minimal Lin28 protein bound to several let-7 precursor RNAs (Nam et al., 2011; Loughlin et al., 2011). Confirming earlier results, these studies revealed that both the CSD and CCHC zinc fingers of Lin28 interact with conserved residues in the preE. Specifically, the Lin28 CSD is inserted into the apical point of the precursor loop, while the CCHC zinc fingers dimerize on a GGAG motif adjacent to the Dicer cleavage site (Nam et al., 2011). A flexible linker peptide between these domains allows for Lin28 to recognize any let-7 isoform, explaining the ability of Lin28 proteins to repress all let-7 family members despite differences in their loop sequence and length. The CCHC zinc fingers partially melt the miRNA stem upon binding, disrupting its double-stranded nature near the base of the

loop and rendering the miRNA refractory to Dicer cleavage (Figure 1.1C). Lin28 represents one of the first of a growing family of RNA-binding proteins that specifically recognizes the terminal loop of miRNAs to control expression (Michlewski et al., 2008; Michlewski et al., 2010; Trabucchi et al., 2009). This novel regulatory process revealed a potentially powerful strategy for cells to regulate specific subset(s) of miRNAs. Indeed, a few non-let-7 miRNAs have been verified as Lin28 substrates after the identification of conserved sequences in their terminal loops, yet the extent of regulation appears quite subtle compared to the repression of let-7 (Heo et al., 2009).

Uncovering the regulatory mechanism of Lin28 towards let-7

The recognition and binding of Lin28 to the preE of let-7 precursors *in vitro* is sufficient to inhibit enzymatic activity of miRNA-processing enzymes, yet the physiological relevance of these *in vitro* observations and how this binding leads to the loss of mature miRNA levels initially remained unresolved. Careful examination of Lin28-associated RNAs revealed a high molecular weight let-7 isoform containing a stretch of 3' untemplated uridines (Heo et al., 2008). This modified pre-let-7 is a poor substrate for Dicer processing, is difficult to detect in total RNA, and has a shortened half-life in cellular extracts, implying that pre-let-7 uridylation is a signal for degradation (Heo et al., 2008). Biochemical purifications of Lin28-containing ribonucleoprotein complexes identified Zcchc11 (also known as TUT4), a non-canonical poly(A) polymerase, as the enzyme responsible for pre-let-7 uridylation (Hagan et al., 2009; Heo et al., 2009). Previously identified as a regulator of Toll-Like Receptor signaling in macrophages and supporting polyuridylation activity *in vitro* (Minoda et al., 2006; Kwak & Wickens, 2007), Zcchc11 was the first Terminal Uridyl Transferase (TUTase) identified to uridylate miRNAs

(Hagan et al., 2009; Heo et al., 2009; Jones et al., 2009). Lin28-binding to pre-let-7 recruits Zcchc11 to catalyze the addition of an oligo-uridine 3' tail *in vitro*, and upon Zcchc11 depletion in ESCs let-7 miRNAs are specifically de-repressed in a manner similar to Lin28 knockdown, albeit more modestly (Hagan et al., 2009; Heo et al., 2009). A possible Zcchc11 ortholog, PUP-2, was implicated in control of let-7 biogenesis in *C. elegans* indicating possible conservation of this regulatory mechanism (Lehrbach et al., 2009). A more recent study however indicated the Lin28-mediated control of let-7 expression in nematode worms does not involve PUP-2 or pre-let-7 uridylation but rather the blockade occurs in the nucleus (van Wynsberghe et al., 2011). Nevertheless in both models Lin28 represses the accumulation of the functional let-7 miRNAs in worm embryos as in mammals.

The catalyzation of a short uridine tail on pre-let-7 as a result of Lin28 binding seemed to commit the miRNA to a degradation pathway, but the major cytoplasmic RNA degradation machinery is not involved in this pathway (Hagan et al., 2009). RNA affinity purification experiments identified the disease gene Dis3l2 (Dis3-like 2) as the exonuclease responsible for the destruction of uridylated pre-let-7. Dis3-related proteins are conserved across eukaryotres and play a central role in RNA stability in the budding yeast *S. cerevisiae* (Staals et al., 2010; Tomecki et al., 2010). This enzyme family—composed of Dis3, Dis3l1, and Dis3l2—was only recently characterized in mammalian cells, yet no endogenous RNA substrates were known before the identification of Dis3l2 in let-7 turnover. Surprisingly, a genome-wide association study (GWAS) found Dis3l2 to be the causal gene in Perlman Syndrome, a rare pediatric disease characterized by fetal overgrowth and high mortality. Patients who survive beyond the first few months of birth face a high probability of renal Wilms tumors and high mortality for the duration of their lives (Astuti et al., 2012). The pathology of Perlman syndrome was reminiscent of

mutations in heterochronic genes that alter development and growth rates, suggesting that Dis3l2 could be a regulator of human developmental timing pathways (Astuti et al., 2012; see below). Chang et al. (2013) determined that Dis3l2 specifically recognizes RNAs with homopolymeric stretches of uridines and preferentially degrades these substrates over non-uridylated RNAs. Knockdown of Dis3l2 in embryonic stem cells led to the accumulation of uridylated pre-let-7, yet mature levels of let-7 miRNAs did not change, presumably because the uridylated miRNA species remained refractory to Dicer cleavage. This study confirmed that Dis3l2, independently found to regulate human growth rates, also plays a role in the Lin28-let-7 pathway, one of the most ancient heterochronic pathways. The identification of Dis3l2 also established a full complement of presumed let-7 regulators, defining a complete pathway from recognition, binding, uridylation, and degradation.

Lin28 is a stem cell pluripotency factor

Lin28 is highly expressed in undifferentiated cells and represents part of a pluripotency network in these cells with the Lin28A promoter region occupied by key ESC transcription factors including Oct4, Sox2 and Nanog (Marson et al., 2008). During embryonic development Lin28 levels decline allowing an accumulation of let-7 miRNAs. let-7 induction inhibits the self-renewal of undifferentiated cells and promotes differentiation. Indeed, introduction of synthetic let-7 is sufficient to rescue the compromised differentiation phenotype of miRNA-deficient ESCs. let-7 directly targets several ‘stemness’ factors including c-Myc, Sall4 and Lin28, all of which have let-7 binding sites in their 3’UTR (Melton et al., 2010). One of the first studies reporting the derivation of human of iPSCs utilized ectopic expression of Oct4, Sox2, Nanog, and Lin28 as reprogramming factors (Yu et al., 2007). This observation, together with the report that

antagonizing let-7 using antisense oligonucleotide inhibitors enhances reprogramming of mouse fibroblasts to iPSCs, indicates that Lin28-mediated inhibition of let-7 expression promotes cell de-differentiation during reprogramming (Melton et al., 2010). Indeed, it was found that Lin28 accelerates reprogramming by promoting cell proliferation (Hanna et al., 2009). Together, these results indicate that the Lin28/let-7 axis operates as a bistable switch functioning to maintain either a differentiated or embryonic cell fate and this pathway can be exploited to manipulate cellular pluripotency.

Lin28 regulates development and translation

Lin28 is dynamically regulated during development but is stably expressed in several embryonic cell types. While the molecular role of Lin28 in the regulation of let-7 expression is now quite well established, studies in various cellular and developmental contexts raise the possibility that Lin28 may have additional regulatory roles and bind certain mRNAs to promote their translation. The relative contribution of Lin28-mediated repression of let-7 and the miRNA-independent functions of Lin28 in most of these contexts remains to be determined. Several reports indicate that Lin28 has let-7 independent roles in the worm heterochronic pathways and in mammalian embryogenesis (Vadla et al., 2012; Xu et al., 2009). Lin28 promotes the rapid proliferation of mouse and human ESCs and directly binds several mRNAs including those encoding certain cell cycle regulatory factors controlling G2/S to M transition, Histone H2a, as well as Oct4 (Xu et al., 2009; Qiu et al., 2009). Indeed, genome-wide analysis of Lin28-associated mRNAs in hESCs revealed an enrichment of metabolic genes and ribosomal protein mRNAs, connecting metabolic regulation to ESC proliferation (Peng et al., 2011). Lin28-mediated translational enhancement requires RNA Helicase A and relies on sequence-specific

binding elements within the coding region of these mRNAs. The mechanism underlying the interaction between Lin28, mRNA, and the translation machinery remains poorly understood (Jin et al., 2011). Lin28 resides in polysomes of muscle progenitors and can enhance the translation of IGF-2 mRNA (Polesskaya et al., 2007). During neurogliogenesis Lin28 levels decrease rapidly upon retinoic acid-induced differentiation of mouse embryonal carcinoma cells (Wu & Belasco, 2005). Forced expression of Lin28 in differentiation conditions allows neural differentiation but inhibits the transition to mature glia. Interestingly, with constitutive expression of Lin28 a significant shift in gene expression pattern occurs before let-7 accumulates, and the overexpression of a Lin28 mutant permitting let-7 accumulation remains inhibitory to glial differentiation, implying let-7-independent functions of Lin28 in glial development (Balzer et al., 2010). Studies of intrinsic neural signaling in the developing mouse brain have linked Lin28 to the specific temporal regulation of a critical subset of neuronal mRNAs (Huang et al., 2012). Brain-Derived Neurotrophic Factor (BDNF) promotes the differentiation of excitatory neurons by selectively promoting the translation of a small subset of genes. Rather than directly promoting translational efficiency, BDNF instead stimulates Lin28 expression which blocks let-7 biogenesis leading to lower levels of let-7 and a corresponding de-repression of let-7 target genes including Dicer. Increased Dicer expression leads to an overall increase in miRNA levels (other than let-7 family members) that further repress their target mRNAs through canonical miRNA silencing activity. Thus, let-7 target mRNAs are elevated and proportionally overrepresented in differentiated neurons following BDNF stimulation (Huang et al., 2012). Through this Lin28-mediated mechanism developing neurons specifically increase the translation of a minor subset of mRNAs in response to BDNF.

In the last few years there have been several other independent studies examining the association of Lin28 with endogenous mRNAs (Wilbert et al., 2012; Cho et al., 2012; Hafner et al., 2013). These studies vary significantly in their findings but establish that Lin28 often binds to motifs containing GGAG residues or other G-rich sequences, and that this interaction has dramatic effects on downstream translation. One study indicates that Lin28 binding disrupts specific splicing patterns (Wilbert et al., 2012), while another suggests that Lin28 binding blocks mRNA degradation and indirectly leads to a subsequent increase in protein levels including that of Lin28 itself (Hafner et al., 2013), while a third finds that Lin28 localizes to the endoplasmic reticulum (ER) and inhibits ER-associated translation and thus secretory activity (Cho et al., 2012). These variable findings are all intriguing for their discoveries and proposed mechanisms, but resolving the clear conflicts amongst these reports must be accomplished before it is understood how Lin28 proteins regulate the activity of mRNAs in various cell types.

Lin28 is also highly expressed in mouse and human primordial germ cells (PGCs). Using an *in vitro* ESC differentiation and RNAi-based strategy, Lin28 was found to be an essential regulator of PGC formation in this system through inhibition of let-7 maturation and consequential induction of the let-7 target gene Blimp1 (West et al., 2009). Lin28 proteins also have an important role in hematopoiesis. Lin28B is abundantly expressed in fetal hematopoietic stem cells (HSCs) while it is absent in terminally differentiated blood cells and adult HSCs. The specific expression of Lin28B in these cells leads to reduced levels of let-7 miRNAs and is responsible for the embryonic-specific lineage commitment of fetal HSC-derived blood cells (Yuan et al., 2012). Overexpression of Lin28A in adult HSCs reprograms these cells to repopulate donor animals with fetal-specific cell types of all lymphoid lineages (Yuan et al., 2012). The ability for Lin28 to enhance the reprogramming of differentiated cells to an

embryonic state, whether iPSCs or fetal HSCs, is further supported by its ability to mediate wound healing in certain contexts (Ramachandran et al., 2010). Unlike mammals, ray-finned fish can repair retinal damage by a program of dedifferentiation to retinal progenitors followed by selective differentiation so as to replace damaged tissues. Upon injury several pluripotency factors including c-Myc and Lin28 are rapidly upregulated in response to the dedifferentiation transcription factor Ascl1. The rapid increase in Lin28 represses let-7 miRNAs, which normally target dedifferentiation genes including c-Myc, Pax6, Ascl1, and Lin28 itself, and therefore Lin28 is required for cell dedifferentiation. The upregulation of Lin28 and subsequent repression of let-7 miRNAs in response to injury provide telling insight into how cells manage their own reprogramming in response to transitional phases and times of stress. The emerging diversity of multiple developmental pathways involving Lin28 shows the importance and breadth of miRNA and mRNA regulation in normal development.

Oncogenic role of Lin28A and Lin28B

After elucidating the Lin28/let-7 connection, several lines of evidence supported the potential role for Lin28 as an oncogene, including the direct targeting of oncogenes by let-7 miRNAs, the embryonic-specific expression of Lin28, and the ability for Lin28 to synergize in somatic cell reprogramming with several known oncogenes (Johnson et al., 2005; Hanna et al., 2009; Mayr et al., 2007; Sampson et al., 2007). Indeed, Lin28 overexpression is sufficient to transform NIH/3T3 cells. Furthermore, depletion of Lin28B in human leukemia cells increases let-7 levels, reduces proliferation, and leads to the decrease of let-7 targets including Myc. Importantly, an array of human hepatocellular carcinoma lines shows an inverse correlation between the expression of Lin28B and all let-7 family members, while high levels of Lin28B are

strongly correlated with early tumor recurrence (Viswanathan et al., 2009). Recent studies have implicated Lin28A or Lin28B upregulation in a growing list of various different cancers and in some cases expression correlates with advanced tumor-stage and poor prognosis. Also an interesting emerging concept is that Lin28A/B expression may characterize distinct tumorigenic subpopulations of cells within a single tumor, giving rise to so-called tumor initiating or cancer stem cells (CSCs) (Zhang et al., 2012). This is intriguing given the parallels of the role of Lin28 in maintaining ESC properties and has important implications for possible therapeutic intervention. Together, these studies underscore the importance of Lin28 proteins in promoting and characterizing various human malignancies.

The clinical relevance of Lin28A/B overexpression is now readily appreciated and recent work is beginning to uncover the mechanism underlying this aberrant expression. Lin28B is potently upregulated in an inducible Myc-mediated lymphoma model and leads to the dramatic reduction of let-7 levels (Chang et al., 2009). Myc binds directly to the Lin28B promoter and drives its expression, while loss of Lin28B in this model disrupts transformation, partly due to an increase in let-7 and subsequent repression of Myc. The establishment, therefore, of a Myc/Lin28B/let-7 oncogenic circuit is necessary and sufficient to transform cells in at least one context. Lin28B expression is also increased by the Src oncoprotein in an inducible breast cancer model (Iliopoulos et al., 2009). Transient activation of Src leads to an inflammatory expression pattern, which ultimately leads to complete cellular transformation. This regulatory switch relies on increased NF-kB signaling; expression of IL-6, a pro-inflammatory cytokine; decreased let-7 levels; and the activation of Lin28B. Intriguingly, IL-6 mRNA is targeted by let-7 miRNAs, and NF-kB binds and enhances expression at the Lin28B promoter. Antagonizing this circuit is sufficient to inhibit transformation and ameliorate the inflammatory gene signature. Similarly,

transient introduction of any of these perturbations is sufficient to drive transformation and establish this feedback loop, which persists for many generations in culture (Iliopoulos et al., 2009). Together these studies provide insight into the direct activation of oncogenic Lin28B and connect miRNA dysregulation to well-known inflammatory and oncogenic expression patterns. How exactly expression of Lin28A is regulated in cancer remains largely unknown but a recent study has unveiled a role for Tristetraprolin (TTP), an RNA binding protein, which recognizes sequence elements in the 3' UTR and facilitates degradation of Lin28A mRNA (Kim et al., 2011). TTP expression is therefore positively correlated with let-7 expression.

Certain types of cancer appear to predominantly express either Lin28A or Lin28B (Thornton & Gregory, 2012). This raises the important question as to the functional equivalence of these paralogous genes. While both proteins selectively repress let-7 expression recent work indicates the mechanism by which this is accomplished is different (Piskounova et al., 2011). The role of the TUTase Zcchc11 in the Lin28 pathway was established in mouse ESCs and embryonal carcinoma cells that predominantly express Lin28A (Hagan et al., 2009; Heo et al., 2009). Since both Lin28A and Lin28B can enhance the activity of the TUTase in biochemical assays *in vitro*, it was surprising that Zcchc11 depletion by shRNA in Lin28B-expressing human cancer cells does not lead to let-7 elevation. Lin28A-expressing human cancer cells are, however, sensitive to Zcchc11 depletion, as knockdown of Zcchc11 increases let-7 levels to a similar extent as observed in ESCs. An explanation for these initially perplexing data came from an analysis of the subcellular localization of Lin28A and Lin28B proteins. Lin28A was detected predominantly in the cell cytoplasm where it interacts with Zcchc11 to repress pre-let-7, while Lin28B resides in the nucleolus and represses pri-let-7 through a TUTase-independent mechanism. Lin28B contains a nuclear localization signal (NLS) and a nucleolar localization

signal (NoLS), each of which is sufficient to localize a fluorescent transgene to its appropriate cellular compartment (Figure 1.1B). Although both proteins act identically in biochemical assays, their non-overlapping localization explains their differential reliance on pre-let-7 uridylation by Zcchc11. In support of this finding Zcchc11 depletion inhibited growth of Lin28A- but not Lin28B-expressing human breast tumors in mouse xenograft assays (Piskounova et al., 2011). Initially thought to be functionally identical and distinguished primarily by expression pattern, the distinct mechanisms of Lin28A and Lin28B in human cancers further refine the view of disease-specific miRNA regulation.

Role of Lin28 proteins in coordinating growth and metabolism

The Lin28/let-7 axis is highly conserved across the animal kingdom in organisms as evolutionarily distant as nematode worms and humans. The first clues that Lin28 family members have retained their heterochronic function in humans came from genome-wide association studies (GWAS). Lin28B was found to be one of several novel loci associated with human height, age of puberty onset, age of menopause, and body-mass index (BMI) (Lettre et al., 2008; Ong et al., 2009; Sulem et al., 2009; He et al., 2009; Perry et al., 2009). A transgenic mouse line overexpressing Lin28A recapitulated several of these developmental timing phenotypes. In these animals low levels of ectopic Lin28A delayed signs of aging including later cessation of growth, later time to first estrus and first litter, and delayed vaginal opening (Zhu et al., 2011). Interestingly, Lin28A and Lin28B transgenic animals are more sensitive to insulin and have reduced peripheral glucose levels. let-7 overexpressing mice have the opposite phenotype of lower insulin sensitivity and higher glucose levels (Zhu et al., 2011a; Frost et al., 2011). The connection between the Lin28/let-7 axis and metabolism seems therefore to be largely controlled

through let-7-dependent regulation. Numerous metabolic genes are direct let-7 targets, notably those in the Insulin-PI3K-mTOR pathway, including INSR, IGF1R, and IRS2 (Zhu et al., 2011a; Frost et al., 2011). The increased size and metabolic differences observed in Lin28A overexpressing animals is rescued by pharmacologically inhibiting the mTOR pathway or overexpressing a let-7 transgene that bypasses Lin28 regulation. Furthermore, meta-analyses of published data confirmed that metabolic genes targeted by let-7 are associated with Type 2 Diabetes (T2D) in humans (Zhu et al., 2011a; Frost et al., 2011). The significant effect of Lin28 and let-7 regulation on metabolism is a surprising one, considering the initial discovery of this pathway in regulating development. However, it poses the intriguing possibility that embryonic development and metabolism are co-regulated over time such that the Lin28/let-7 axis functions early in development to integrate cell proliferation with the high metabolic demands of the cells in a rapidly growing embryo, while in the adult, the pathway coordinates systemic metabolism at the organismal level (Figure 1.2) (Reviewed in Shyh-Chang & Daley, 2013).

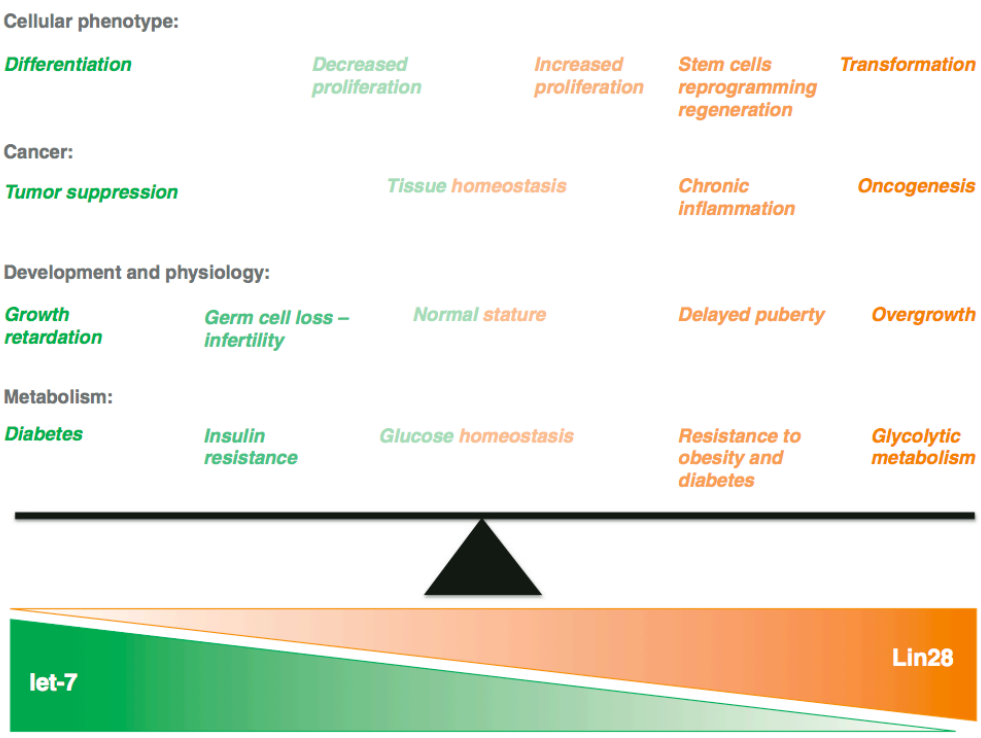


Figure 1.2: Balancing the Lin28/let-7 axis in development and disease

The future of Lin28 and let-7

The regulation of let-7 miRNAs by Lin28 is a rapidly growing field and points to the importance of small RNA metabolism in disparate realms of mammalian biology. Ongoing studies will develop our understanding of the molecular and cellular mechanisms of this pathway. A more complete understanding of how Lin28A and Lin28B are regulated and how these proteins help maintain precise levels of let-7 miRNAs in the context of development, tissue homeostasis/regeneration, aging, and oncogenesis will facilitate the development of ways to exploit this regulatory pathway by manipulating the Lin28/let-7 axis for novel treatments of human diseases including cancer and diabetes. One complication arising from the understanding of Lin28/let-7 regulation is the dual effect of Lin28 in being both protective and causal in different types of human disease. The animal models of Lin28A/B overexpression indicate the protective effect let-7 inhibition has in T2D development (Zhu et al., 2011; Zhu et al., 2011a; Frost et al., 2011). However, Lin28 functions as an oncogene in cancer raising the likelihood that persistent Lin28 expression (or let-7 repression) will likely lead to hyperplastic disorders and in certain cases cancer. However, a recent proof-of-principle is encouraging in this regard; inhibition of let-7 with an antimiR improved insulin sensitivity and glucose tolerance, raising the intriguing possibility that let-7 antimiR treatment could provide a new approach to treating T2D. It is also noteworthy that no cancer developed in antimiR-treated mice, although the 10-week time course was relatively short (Frost et al., 2011). Nevertheless, evidence supports that Lin28 expression levels must be tightly regulated both during developmental stage and cellular context. Indeed, the GWAS findings linking Lin28B to human development underscore this point. Although it may be useful to repress let-7 levels in inflammatory circumstances, doing so to a significant extent may metabolically reprogram cells to an embryonic-like state and lead to the

initiation of cancer. This so-called Warburg Effect (Koppenol et al., 2011) where a switch occurs to glycolytic metabolism, will increase glucose consumption in peripheral tissues, which could protect against diabetes, but this metabolic transformation is a hallmark of cancer.

Fine-tuning of gene expression by miRNAs is known to have significant clinical and biological impacts, but now it seems that subtle modulation of miRNA-regulating factors may have similarly dramatic effects. The elucidation of Lin28 and let-7 in mammalian development and human disease has uncovered many exciting new aspects in various realms of biology. Bridging the gap from embryonic development to metabolism and how these pathways can activate or ameliorate chronic inflammation and cancer, these studies have done a great deal in tying together formerly disparate fields (Figure 1.2). Further work on uncovering mechanistic details of this pathway, especially finding inhibitors and identifying novel regulators, will go a long way in understanding how best to treat and control human diseases.

PAPs and TUTases

The post-transcriptional regulation of miRNAs takes many forms, including the developmental and disease-related regulation of let-7 miRNAs by Lin28 described above. Work published in only the last five years has expanded on some of these modifications and has uncovered an exciting new role for non-templated nucleotide addition in miRNA processing and function. This process occurs broadly across many miRNA substrates and is mostly carried out by a host of enzymes belonging to a specific subfamily of RNA polymerases called Terminal Uridyl Transferases [TUTases]/Poly(U) polymerases [PUPs]. Poly(A) Polymerases [PAPs] also have a role in miRNA modification, potentially as antagonists to other terminal nucleotide modifying enzymes. The pace with which this field has developed and the breadth of activities

described thus far leave many outstanding questions in the field, as a unifying understanding of miRNA nucleotide addition remains elusive.

Addition of monomeric residues to the free ends of polynucleotides has been understood as a central role in biology for decades (Watson & Crick, 1953). The enzymes responsible for this activity in eukaryotes belong to several large superfamilies including those catalyzing DNA replication, DNA repair, transcription, mRNA polyadenylation, tRNA terminal CCA addition, and terminal uridine polymerization (Aravind & Koonin, 2000; Kwak & Wickens, 2007; Rissland et al., 2007). A large subclass of polymerases closely related to eukaryotic DNA-polymerase is the so-called Polymerase-β-like (Polβ) members. Those that specifically modify RNA include mRNA poly(A) polymerases [PAPs], tRNA CCA adding enzymes, and a subgroup of polymerases related to PAPs that have specificity for other nucleotides and may or may not target mRNA (now known as non-canonical poly(A) polymerases). (Holm & Sander, 1995; Yue et al., 1996; Rissland et al., 2007). Together the Polβ RNA polymerases perform a diverse set of functions, many of which remain unappreciated but are increasingly understood as central to eukaryotic gene regulation.

TUTases were first described in the parasite *Trypanosome bruceii* as essential regulators of RNA processing in the mitochondria and the cytoplasm. These TUTases have been heavily studied, and solution structures are described for at least two of these TUTases: MEAT1, a mitochondrial RNA editing TUTase; and RET2, a related TUTase with similar activity. Despite disparate enzymatic activities ranging from rRNA-specific uridine addition to uridine insertional mutagenesis, the active sites of trypanosome TUTases are remarkably similar (Aphasizheva et al., 2004; Stagno et al., 2007; Stagno et al., 2007a; Stagno et al., 2010). The solution structure of MEAT1, a mitochondrial *T. bruceii* TUTase revealed that the specific selection of uridine

nucleoside triphosphates occurs in the absence of a substrate RNA, and has provided the first evidence as to how TUTases use critical residues in their active sites to select and incorporate nucleotides. TUTase activity requires the divalent cations Mg^{2+} or Mn^{2+}, and forms a network between water molecules and several conserved amino acid side chains near the active site. Interestingly, MEAT1 also specifies uridine incorporation by using amino acids that are poorly conserved among other TUTases, implying that various TUTases can be classified by their active site structures, but may use divergent residues for similar functions. It is predicted that minor variations in amino acid residues near the active site may alter nucleotide and substrate specificity, perhaps explaining, for example, why *Tb*TUT4 permits cytidine transferase activity, while most other *Tb*TUTases exclusively utilize uridine (Stagno et al., 2010). Underscoring the complexity of TUTase substrate detection and selectivity, a crystal structure of *S. pombe* Cid1 was recently reported, finding specific differences between the yeast and parasite enzymes (Yates et al., 2012). While in *Trypanosomes* uridine selectivity requires organized water molecules and intimate interactions with serine and lysine residues, *S. pombe* and mammalian TUTases contain a conserved histidine residue that makes direct contact with the aromatic ring of free UTP. This single residue is sufficient to account for uridine selectivity, as mutating this residue to an alanine converts the TUTase to a PAP with nearly exclusive preference for free ATP. Furthermore, a cryptic basic region mediates sequence binding and appears to feed bound RNA towards the active site where bound UTP resides. These two novel and unique mechanisms set mammalian and fission yeast TUTases apart from their bacterial and parasite counterparts, and may explain their most basic form of nucleotide and substrate specificity.

TUTase Identification and Roles in Gene Regulation

The firsts identification of mammalian Poly(U) Polymerase activity only occurred in the late 1990's with the identification of a TUTase that specifically uridylated the U6 snRNA to facilitate splicing turnover (Trippe et al., 1998; Trippe et al., 2003; Trippe et al., 2006). Characterization of this enzyme revealed specific activity towards the U6 snRNA containing one, two, or three terminal uridines with the longest RNA species as the preferred substrate. With the rising interest in non-canonical poly(A) polymerases, a surprising study found that a class of enzymes related to U6 TUTase/TUT1/TUTase6/PAPD2/STARPAP but predicted to carry out poly(A) activity were indeed TUTases as well (Kwak & Wickens; 2007). While searching for homologs of a yeast PAP across diverse evolutionary taxa, this group found that proteins encoded in organisms from *A. thaliana* to humans shared a significant degree of homology with PAPs, but preferentially incorporated uridine residues. Almost simultaneously another group found that *S. pombe* Cid1, a homolog of the yeast Cid1 Poly(A) Polymerase, had a strong preference for uridines and that the human homolog Hs2/Zcchc6/TUTase 7 carried out similar activity (Rissland et al., 2007). Soon thereafter it was found in fission yeast that mRNA uridylation promotes decapping and mRNA turnover, suggesting that this conserved mechanism of 3' poly(U) modification may be an ancient form of gene regulation in diverse species (Rissland & Norbury, 2009). The catalog of mammalian TUTases stands at seven distinct members, several of which share a significant degree of homology. The piecemeal identification of different family members has left TUTase nomenclature convoluted and without an accepted naming system (Figure 1.3), yet the progress made in the last several years has brought to light the diversity of roles that these enzymes can play.

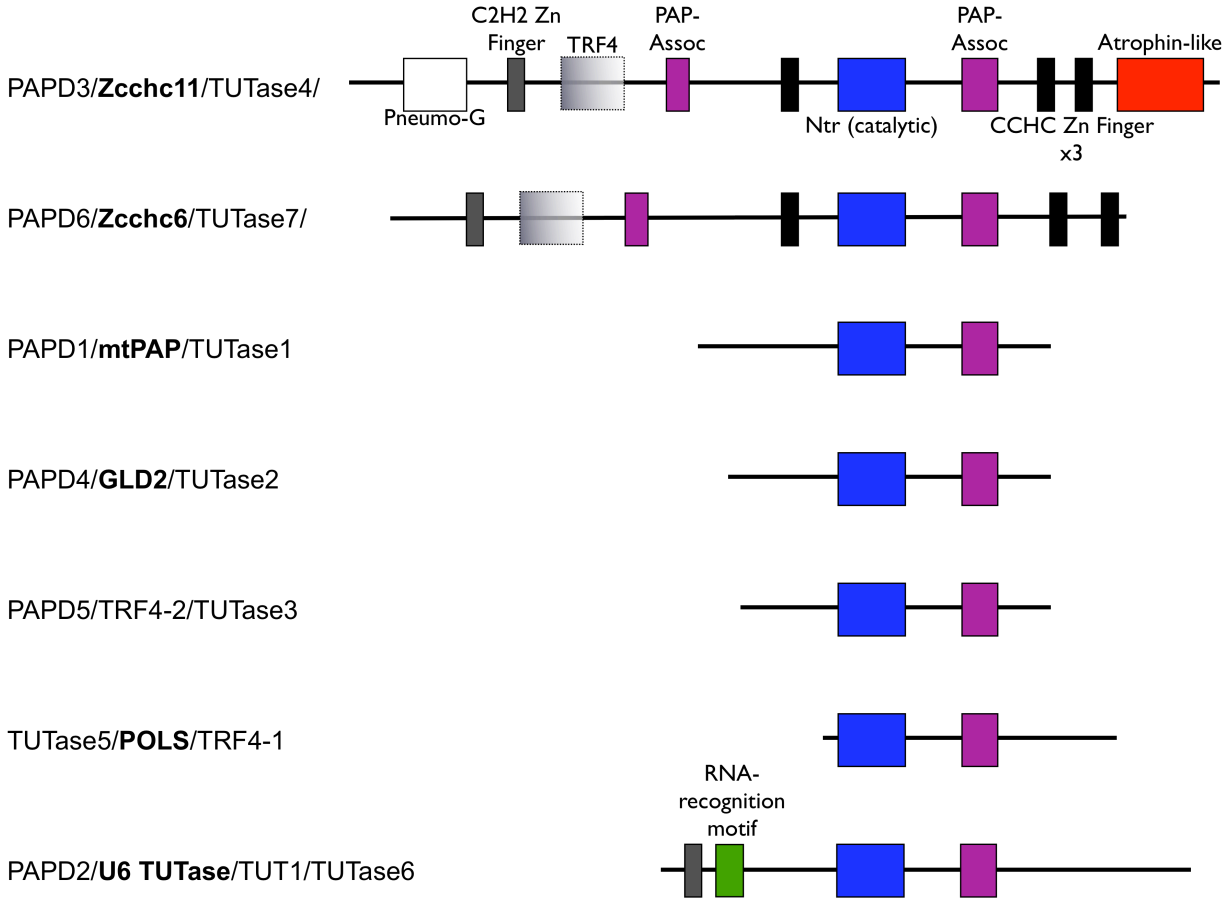


Figure 1.3: Diagram of Mammalian TUTases

Preferred names shown in bold. All colored domains correspond to the labels as shown. Adapted from Heo et al., 2009.

One of the few classes of protein-coding genes that lack a poly(A) tail is the replicative histone family. Rather than receiving a canonical 3' end, many histone transcripts encode a 3' stem loop structure that forms a stable hairpin as the transcript terminates. This loop is necessary and sufficient for histone mRNA regulation during the cell cycle, where histone protein levels must double during every round of DNA replication and histone mRNA levels must be rapidly silenced at the end of S phase so as to maintain a stoichiometric balance between DNA and histones. This terminal loop functions by recruiting Stem Loop Binding Protein (SLBP), which conveys all regulatory information for histone turnover during the cell cycle (Dominski & Marzluff, 1999). Examining histone mRNA termini after inducing degradation by hydroxyurea (HU) treatment revealed widespread oligouridylation of several replicative histones transcripts. This modification induced 5' decapping and simultaneous degradation from both the 5' and 3' ends. (Mullen & Marzluff, 2008). Uridylation-mediated degradation of histone mRNAs is at least partially dependent on the LSM1-7 complex, which forms a heptameric ring and binds to both U-rich mRNAs and SLBP. This study identified the exosome and Xrn1 as mediators of 3'-5' and 5'-3' degradation, respectively, and the two TUTases PAPD1 and PAPD5 as the enzymes responsible for 3' uridylation. Later work found another RNA exonuclease responsible for histone turnover, namely Eri1/3'hEXO, with almost no role for the exosome. Furthermore, it was later determined that Zcchc11 is responsible for cell-cycle-dependent histone mRNA uridylation and that PAPD1 and PAPD5 have other roles including mitochondrial RNA uridylation and nuclear mRNA adenylation (Yang et al., 2006; Hoefig et al., 2013; Schmidt et al., 2010). Clearly the unique instance of histone mRNA turnover is complicated by non-canonical polymerases and nucleases, and several rounds of revision have been made to this dynamic process. The

importance of non-canonical poly(A) polymerase activity, however, is underscored by their essential role in such a highly-conserved and evolutionarily important pathway.

Nucleotidal Transferases and miRNAs

The realization of widespread uridylation and adenylation presented a possibility that non-canonical poly(A) polymerases play a larger role in regulating other RNA substrates. A surprising role for this enzyme family was uncovered in 2009 when Katoh and colleagues (2009) found through mass-spectrometric approaches that the liver-enriched miRNA miR-122 existed as several isoforms varying at the 3' end. Aside from Dicer cleavage variants, one species was found that contained a single 3' adenosine residue that could only have arisen from non-templated addition. A candidate approach led the authors to GLD-2 as the enzyme responsible for this modification, and mice lacking GLD-2—a known non-canonical PAP—showed a complete absence of miR-122 in their livers, indicating that GLD-2-mediated monoadenylation of miR-122 is required for specific miRNA stabilization. A recent publication demonstrated that GLD-2 adenylates and stabilizes several other mature miRNAs aside from miR-122, including several members of the let-7 family (D'ambrogio et al., 2012). GLD-2-mediated uridylation requires no other accessory factors, and is dependent on specific residues in the mature miRNA 3' end. Two related reports underscored the complexity of mature miRNA modifications by non-canonical poly(A) polymerases, in some cases disputing the findings of previous studies of GLD-2. Burroughs et al (2010) found GLD-2 depletion to broadly affect mature miRNA adenylation *without* altering miRNA levels on a genome-wide scale. mRNAs targeted by adenylated miRNAs were derepressed, suggesting that adenylation antagonizes miRNA function. To this point the authors found that miRNAs loaded into Ago2 or Ago3 but not Ago1 had not undergone

3’ adenylation, implying an exclusionary mechanism on the part of Ago proteins for adenylated miRNAs. A related paper found the depletion of each member of the non-canonical poly(A) polymerase family to result in changes at the 3’ ends of various miRNAs. These changes could be the loss of any individual non-templated nucleotide (e.g. adenosine residues in the case of GLD-2) or the loss of nucleotide combinations (e.g. miR-1246 lost non-templated GA couplets after PAPD5/TRF4-2/TUTase3 depletion) (Wyman et al., 2011). Although GLD-2 adenylation activity was the first example of a PAP modifying mature miRNAs, its suspected roles are diverse and often in conflict. The multiple roles non-canonical poly(A) polymerases play in miRNA biogenesis are still rapidly being uncovered and the subtle impacts of their activity may complicate analyses of their activity in mammalian organisms.

Within one month of the description of the GLD-2/miR-122 pathway, a different mature miRNA was found to be targeted by another member of the non-canonical poly(A) polymerase family. A group investigating the GLD-2-like enzyme Zcchc11/TUTase4/Hs3/PAPD3 was able to nicely show that the enzyme had uridine-specific activity towards a small number of mature miRNAs. Upon Zcchc11 depletion there was a potent reduction in a group of pro-inflammatory cytokines including IL-6, TNFα, VEGF, and RANTES, and after examining the 3’ ends of these mRNAs there were, surprisingly, no non-templated uridine residues. Instead, miR-26a/b, a miRNA family targeting IL-6 underwent 3’ monouridylation, which reduced their ability to repress target genes. Synthetic miR-26 carrying U tails of varying lengths lost its potency in a dose-dependent manner, implying that non-templated uridylation disrupts miRNA-mediated repression. Interestingly and in contrast to the effects of GLD-2 on miR-122, Zcchc11-mediated uridylation of miR-26 did not lead to a reduction in miRNA levels but rather seemed to disrupt the function of the pathway through a mechanism that remains unclear (Jones et al., 2009).

Zcchc11 was previously known to mediate the pro-inflammatory cascade from toll-like receptor signaling and was recently shown to disrupt cell cycle regulators to promote proliferation. Interestingly, neither of these activities relied on catalytic residues and instead they were mediated through an N-terminal region of unknown function (Minoda et al., 2006; Blahna et al., 2011).

During the uncovering of this broad class of non-canonical poly(A) polymerases there was a parallel track of research investigating the mechanism of Lin28-mediated let-7 repression as described previously. After the realization that Lin28-bound pre-let-7 underwent non-templated oligouridylation that led to its destruction, there was a great deal of effort to uncover the enzymes responsible for this activity. Our group, along with V. Narry Kim's lab, found that Zcchc11 is also the enzyme responsible for uridylating pre-let-7 in embryonic stem cells and that this activity was specific to let-7 family members, as nearly no other miRNAs were dysregulated upon Zcchc11 depletion (Hagan et al., 2009; Heo et al., 2008; Heo et al., 2009). Upon deeper analysis of miRNA precursors during differentiation, one group found that while Lin28 and Zcchc11 do mediate the oligouridylation of let-7 precursors in embryonic cells, monouridylation by unknown TUTases is widespread across numerous miRNA precursors, and is enriched in adult cells where Lin28 and Lin28B are absent (Newman et al., 2011). The authors of this study also found adenosine and cytidine modifications, but at a much lower level. The oligouridylation mediated by Lin28 drives the destruction of let-7 precursors by Dis3L2, but monomeric nucleotide addition did not appear to mediate the destruction of any miRNA precursors studied. This work was expanded when let-7 precursors were analyzed in Hela cells, which lack Lin28 and Lin28B. In this cellular context, a subset of let-7 precursors is processed by the Microprocessor one nucleotide shorter than is ideal for Dicer processing. In this circumstance a

trio of TUTases interacts with the improper 3' end and adds a single uridine residue that facilitates proper Dicer processing. These three TUTases, Zcchc11, Zcchc6, and GLD-2, all belong to the family of non-canonical poly(A) polymerases, but this common role was unexpected given the specificity of GLD-2 for adenosine, and the previously known role of Zcchc11 in the destruction of the same let-7 precursors. Still, in the absence of these three TUTases this subset of let-7 precursors and several other miRNAs (termed Type II miRNAs) are poorly processed, while other miRNA precursors that contain a proper 2nt 3' overhang (Type I miRNAs) do not rely on monouridylation for their proper processing. Evolutionarily the Type II let-7 members are expressed to a much greater extent than Type I family members implying that monouridylation of let-7 precursors is essential (Heo et al., 2012). The diverse ways in which miRNAs can be modified at various steps along their processing pathway further emphasizes the importance of nucleotide-based alterations in gene expression.

Recently, Zcchc11 knockout mice were created to study the impact on germ-line loss of this TUTase (Jones et al., 2012). Mice lacking Zcchc11 were born at Mendelian ratios and survived for several days after birth, only to die soon thereafter. The few Zcchc11-/- animals that did survive weighed less than their wild-type littermates and expressed uridylated miRNAs to a much lower extent. Interestingly, many miRNAs that underwent a reduction in 3' uridylation are those that target the 3'UTRs of growth factor mRNAs including IGF-1. Zcchc11-/- mice expressed lower levels of serum IGF-1 compared to their wild-type littermates, which likely explains their smaller size and high post-natal mortality. This discovery that Zcchc11-mediated uridylation of mature miRNAs occurs on disparate substrates that together have a common network effect may nicely explain how non-templated uridylation alters specific facets of gene expression without broadly dysregulating PTGS. Although in some instances substrate targeting

seems to be very specific and quite potent, the specificity factors that guide TUTases and PAPs to their substrates remain elusive.

REFERENCES

Ambros, V. and H.R. Horvitz, Heterochronic mutants of the nematode Caenorhabditis elegans. Science, 1984. 226(4673): p. 409-16.

Aphasizheva et al. RNA-editing terminal uridylyl transferase 1: identification of functional domains by mutational analysis. J Biol Chem (2004) vol. 279 (23) pp. 24123-30

Aravind and Koonin. DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. Nucleic Acids Research (1999) vol. 27 (7) pp. 1609-18

Astuti et al. Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. Nature Genetics, 2012. 44(3): p. 277-284.

Balzer, E., et al., LIN28 alters cell fate succession and acts independently of the let-7 microRNA during neurogliogenesis in vitro. Development, 2010. 137(6): p. 891-900.

Bernstein et al. Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature (2001) vol. 409 (6818) pp. 363-6

Blahna et al. Terminal uridyltransferase enzyme zcchc11 promotes cell proliferation independent of its uridyltransferase activity. J Biol Chem (2011) vol. 286 (49) pp. 42381-9

Bohmert et al. AGO1 defines a novel locus of Arabidopsis controlling leaf development. EMBO J (1998) vol. 17 (1) pp. 170-80

Brannan et al. The product of the H19 gene may function as an RNA. Mol Cell Biol (1990) vol. 10 (1) pp. 28-36

Britten and Davidson. Gene regulation for higher cells: a theory. Science (1969) vol. 165 (3891) pp. 349-57

Brockdorff et al. The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. Cell (1992) vol. 71 (3) pp. 515-26

Burroughs et al. A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. Genome Res (2010) vol. 20 (10) pp. 1398-410

Catalanotto et al. Gene silencing in worms and fungi. Nature (2000) vol. 404 (6775) pp. 245

Chang et al. A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. Nature, 2013. 497(7448): p. 244-248

Chang, T.C., et al., Lin-28B transactivation is necessary for Myc-mediated let-7 repression and

proliferation. Proc Natl Acad Sci U S A, 2009. 106(9): p. 3384-9.

Chekulaeva, M. and W. Filipowicz, Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. Curr Opin Cell Biol, 2009. 21(3): p. 452-60.

Cheloufi et al. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. Nature (2010) pp.

Chendrimada et al. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature (2005) vol. 436 (7051) pp. 740-4

Cho et al. LIN28A is a suppressor of ER-associated translation in embryonic stem cells. Cell, 2012. 151(4): p. 765-77.

Cifuentes et al. A Novel miRNA Processing Pathway Independent of Dicer Requires Argonaute2 Catalytic Activity. Science (2010) pp.

F. Crick, Central dogma of molecular biology, Nature 227, 561–563 (1970).

D'ambrogio et al. Specific miRNA Stabilization by Gld2-Catalyzed Monoadenylation. CellReports (2033) vol. 2 (6) pp. 1537-1545

Dominski and Marzluff. Formation of the 3' end of histone mRNA. Gene (1999) vol. 239 (1) pp. 1-14

Elkayam et al. The structure of human argonaute-2 in complex with miR-20a. Cell (2012) vol. 150 (1) pp. 100-10

Fire et al. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature (1998) vol. 391 (6669) pp. 806-11

Frost, R.J. and E.N. Olson, Control of glucose homeostasis and insulin sensitivity by the Let-7 family of microRNAs. Proc Natl Acad Sci USA, 2011. 108(52): p. 21075-80.

Gregory et al. Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. Cell (2005) vol. 123 (4) pp. 631-40

Hafner et al. Identification of mRNAs bound and regulated by human LIN28 proteins and molecular requirements for RNA recognition. RNA, 2013. 19: p. 613-626.

Hagan, J.P., E. Piskounova, and R.I. Gregory, Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nat Struct Mol Biol, 2009. 16(10): p. 1021-5.

Hamilton and Baulcombe. A species of small antisense RNA in posttranscriptional gene silencing in plants. Science (1999) vol. 286 (5441) pp. 950-2

Hammond et al. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature (2000) vol. 404 (6775) pp. 293-6

Hammond et al. Argonaute2, a link between genetic and biochemical analyses of RNAi. Science (2001) vol. 293 (5532) pp. 1146-50

Hanna, J., et al., Direct cell reprogramming is a stochastic process amenable to acceleration. Nature, 2009. 462(7273): p. 595-601.

He, C., et al., Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. Nat Genet, 2009. 41(6): p. 724-8.

Heo et al. Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. Cell (2012) vol. 151 (3) pp. 521-32

Heo, I., et al., Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol Cell, 2008. 32(2): p. 276-84.

Heo, I., et al., TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell, 2009. 138(4): p. 696-708.

Hoefig et al. Eri1 degrades the stem-loop of oligouridylated histone mRNAs to induce replication-dependent decay. Nature Structural & Molecular Biology (2013) vol. 20 (1) pp. 73-81

Holm and Sander. DNA polymerase beta belongs to an ancient nucleotidyltransferase superfamily. Trends Biochem Sci (1995) vol. 20 (9) pp. 345-7

Huang, Y.W., et al., Dual regulation of miRNA biogenesis generates target specificity in neurotrophin-induced protein synthesis. Cell, 2012. 148(5): p. 933-46.

Hutvágner and Zamore. A microRNA in a multiple-turnover RNAi enzyme complex. Science (2002) vol. 297 (5589) pp. 2056-60

Hutvágner et al. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science (2001) vol. 293 (5531) pp. 834-8

Iliopoulos, D., H.A. Hirsch, and K. Struhl, An epigenetic switch involving NF-kappaB, Lin28, Let-7 MicroRNA, and IL6 links inflammation to cell transformation. Cell, 2009. 139(4): p. 693-706.

Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* 3, 318–356 (1961).

Jacquier. The complex eukaryotic transcriptome: unexpected pervasive transcription and novel small RNAs. Nat Rev Genet (2009) vol. 10 (12) pp. 833-44

Jin, J., et al., Evidence that Lin28 stimulates translation by recruiting RNA helicase A to polysomes. Nucleic Acids Res, 2011. 39(9): p. 3724-34.

Johnson, S.M., et al., RAS is regulated by the let-7 microRNA family. Cell, 2005. 120(5): p. 635-47.

Jones et al. Zcchc11-dependent uridylation of microRNA directs cytokine expression. Nat Cell Biol (2009) pp.

Jones, M.R., et al., Zcchc11-dependent uridylation of microRNA directs cytokine expression. Nat Cell Biol, 2009. 11(9): p. 1157-63.

Kanellopoulou et al. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes & Development (2005) vol. 19 (4) pp. 489-501

Katoh et al. Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. Genes & Development (2009) vol. 23 (4) pp. 433-8

Kawamata, T. and Y. Tomari, Making RISC. Trends Biochem Sci, 2010. 35(7): p. 368-76.

Khvorova et al. Functional siRNAs and miRNAs exhibit strand bias. Cell (2003) vol. 115 (2) pp. 209-16

Kim et al. MicroRNA-directed transcriptional gene silencing in mammalian cells. Proc Natl Acad Sci USA (2008) vol. 105 (42) pp. 16230-5

Kim, C.W., et al., Ectopic over-expression of tristetraprolin in human cancer cells promotes biogenesis of let-7 by down-regulation of Lin28. Nucleic Acids Res, 2011.

Koppenol, W.H., P.L. Bounds, and C.V. Dang, Otto Warburg's contributions to current concepts of cancer metabolism. Nat Rev Cancer, 2011. 11(5): p. 325-37.

Krol, J., I. Loedige, and W. Filipowicz, The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet, 2010. 11(9): p. 597-610.

Kwak, J.E. and M. Wickens, A family of poly(U) polymerases. RNA, 2007. 13(6): p. 860-7.

Lagos-Quintana et al. Identification of novel genes coding for small expressed RNAs. Science (2001) vol. 294 (5543) pp. 853-8

Lau et al. An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science (2001) vol. 294 (5543) pp. 858-62

Lee and Ambros. An extensive class of small RNAs in Caenorhabditis elegans. Science (2001) vol. 294 (5543) pp. 862-4

Lee et al. MicroRNA maturation: stepwise processing and subcellular localization. EMBO J (2002) vol. 21 (17) pp. 4663-70

Lee et al. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell (1993) vol. 75 (5) pp. 843-54

Lee et al. The nuclear RNase III Drosha initiates microRNA processing. Nature (2003) vol. 425 (6956) pp. 415-9

Lee, R.C., R.L. Feinbaum, and V. Ambros, The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell, 1993. 75(5): p. 843-54.

Lehrbach, N.J., et al., LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans. Nat Struct Mol Biol, 2009. 16(10): p. 1016-20.

Lettre, G., et al., Identification of ten loci associated with height highlights new biological pathways in human growth. Nat Genet, 2008. 40(5): p. 584-91.

Lewis et al. Prediction of mammalian microRNA targets. Cell (2003) vol. 115 (7) pp. 787-98

Liu et al. R2D2, a bridge between the initiation and effector steps of the Drosophila RNAi pathway. Science (2003) vol. 301 (5641) pp. 1921-5

Liu et al. Argonaute2 is the catalytic engine of mammalian RNAi. Science (2004) vol. 305 (5689) pp. 1437-41

Loughlin, F.E., et al., Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28. Nat Struct Mol Biol, 2011.

Lu et al. MicroRNA expression profiles classify human cancers. Nature (2005) vol. 435 (7043) pp. 834-8

Ma et al. Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain. Nature (2004) vol. 429 (6989) pp. 318-22

MacRae et al. In vitro reconstitution of the human RISC-loading complex. Proc Natl Acad Sci USA (2008) vol. 105 (2) pp. 512-7

Marson, A., et al., Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell, 2008. 134(3): p. 521-33.

Martinez and Gregory. Argonaute2 expression is post-transcriptionally coupled to microRNA abundance. RNA (2013) pp.

Mayr, C., M.T. Hemann, and D.P. Bartel, Disrupting the pairing between let-7 and Hmga2

enhances oncogenic transformation. Science, 2007. 315(5818): p. 1576-9.

Melton, C., R.L. Judson, and R. Blelloch, Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. Nature, 2010. 463(7281): p. 621-6.

Michlewski, G. and J.F. Caceres, Antagonistic role of hnRNP A1 and KSRP in the regulation of let-7a biogenesis. Nat Struct Mol Biol, 2010. 17(8): p. 1011-8.

Michlewski, G., et al., Posttranscriptional regulation of miRNAs harboring conserved terminal loops. Mol Cell, 2008. 32(3): p. 383-93.

Minoda, Y., et al., A novel Zinc finger protein, ZCCHC11, interacts with TIFA and modulates TLR signaling. Biochem Biophys Res Commun, 2006. 344(3): p. 1023-30.

Moss, E.G., R.C. Lee, and V. Ambros, The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA. Cell, 1997. 88(5): p. 637-46.

Mullen and Marzluff. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. Genes & Development (2008) vol. 22 (1) pp. 50-65

Murchison et al. Characterization of Dicer-deficient murine embryonic stem cells. Proc Natl Acad Sci USA (2005) vol. 102 (34) pp. 12135-40

Nam, Y., et al., Molecular Basis for Interaction of let-7 MicroRNAs with Lin28. Cell, 2011. 147(5): p. 1080-91.

Newman et al. Deep sequencing of microRNA precursors reveals extensive 3' end modification. RNA (New York, NY) (2011) pp.

Newman, M.A., J.M. Thomson, and S.M. Hammond, Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. RNA, 2008. 14(8): p. 1539-49.

Ong, K.K., et al., Genetic variation in LIN28B is associated with the timing of puberty. Nat Genet, 2009.

Park et al. Dicer recognizes the 5' end of RNA for efficient and accurate processing. Nature (2011) vol. 475 (7355) pp. 201-5

Pasquinelli, A.E., et al., Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature, 2000. 408(6808): p. 86-9.

Peng, S., et al., Genome-wide studies reveal that Lin28 enhances the translation of genes important for growth and survival of human embryonic stem cells. Stem Cells, 2011. 29(3): p. 496-504.

Perry, J.R., et al., Meta-analysis of genome-wide association data identifies two loci influencing age at menarche. Nat Genet, 2009. 41(6): p. 648-50.

Piskounova, E., et al., Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28. J Biol Chem, 2008. 283(31): p. 21310-4.

Piskounova, E., et al., Lin28A and Lin28B Inhibit let-7 MicroRNA Biogenesis by Distinct Mechanisms. Cell, 2011. 147(5): p. 1066-79.

Polesskaya, A., et al., Lin-28 binds IGF-2 mRNA and participates in skeletal myogenesis by increasing translation efficiency. Genes Dev, 2007. 21(9): p. 1125-38.

Qiu, C., et al., Lin28-mediated post-transcriptional regulation of Oct4 expression in human embryonic stem cells. Nucleic Acids Res. 38(4): p. 1240-8.

Ramachandran, R., B.V. Fausett, and D. Goldman, Ascl1a regulates Muller glia dedifferentiation and retinal regeneration through a Lin-28-dependent, let-7 microRNA signalling pathway. Nat Cell Biol, 2010. 12(11): p. 1101-7.

Ratcliff et al. A similarity between viral defense and gene silencing in plants. Science (1997) vol. 276 (5318) pp. 1558-60

Reinhart, B.J., et al., The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature, 2000. 403(6772): p. 901-6.

Rissland et al. Efficient RNA polyuridylation by noncanonical poly(A) polymerases. Molecular and Cellular Biology (2007) vol. 27 (10) pp. 3612-24

Rissland and Norbury. Decapping is preceded by 3' uridylation in a novel pathway of bulk mRNA turnover. Nature Structural & Molecular Biology (2009) vol. 16 (6) pp. 616-23

Rivas et al. Purified Argonaute2 and an siRNA form recombinant human RISC. Nature Structural & Molecular Biology (2005) vol. 12 (4) pp. 340-9

Rybak, A., et al., A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. Nat Cell Biol, 2008. 10(8): p. 987-93.

Sampson, V.B., et al., MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt lymphoma cells. Cancer Res, 2007. 67(20): p. 9762-70.

Schirle and MacRae. The crystal structure of human Argonaute2. Science (2012) vol. 336 (6084) pp. 1037-40

Schmidt et al. The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. RNA (New York, NY) (2010) pp.

Schwarz et al. Asymmetry in the assembly of the RNAi enzyme complex. Cell (2003) vol. 115 (2) pp. 199-208

Shyh-Chang and Daley. Lin28: primal regulator of growth and metabolism in stem cells. Cell Stem Cell (2013) vol. 12 (4) pp. 395-406

Soifer et al. A role for the Dicer helicase domain in the processing of thermodynamically unstable hairpin RNAs. Nucleic Acids Research (2008) vol. 36 (20) pp. 6511-22

Staals et al. Dis3-like 1: a novel exoribonuclease associated with the human exosome. The EMBO Journal, 2010. 29(14): p. 2358-67

Stadtfeld, M. and K. Hochedlinger, Induced pluripotency: history, mechanisms, and applications. Genes Dev, 2010. 24(20): p. 2239-63.

Stagno et al. Dual role of the RNA substrate in selectivity and catalysis by terminal uridylyl transferases. Proc Natl Acad Sci USA (2007) vol. 104 (37) pp. 14634-9

Stagno et al. Structure of the mitochondrial editosome-like complex associated TUTase 1 reveals divergent mechanisms of UTP selection and domain organization. J Mol Biol (2010) vol. 399 (3) pp. 464-75

Stagno et al. UTP-bound and Apo structures of a minimal RNA uridylyltransferase. J Mol Biol (2007) vol. 366 (3) pp. 882-99

Suh, M.R., et al., Human embryonic stem cells express a unique set of microRNAs. Dev Biol, 2004. 270(2): p. 488-98.

Sulem, P., et al., Genome-wide association study identifies sequence variants on 6q21 associated with age at menarche. Nat Genet, 2009. 41(6): p. 734-8.

Tabara et al. The rde-1 gene, RNA interference, and transposon silencing in C. elegans. Cell (1999) vol. 99 (2) pp. 123-32

Takahashi, K. and S. Yamanaka, Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell, 2006. 126(4): p. 663-76.

Thomson, J.M., et al., Extensive post-transcriptional regulation of microRNAs and its implications for cancer. Genes Dev, 2006. 20(16): p. 2202-7.

Thornton, J.E., and Gregory, R.I. (2012). How does Lin28 let-7 control development and disease? Trends Cell Biol *22*, 474-482.

Tomari et al. A protein sensor for siRNA asymmetry. Science (2004) vol. 306 (5700) pp. 1377-80

Tomecki et al. The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. The EMBO Journal, 2010. 29(14): p. 2342-57

Trabucchi, M., et al., The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. Nature, 2009. 459(7249): p. 1010-4.

Trippe et al. A highly specific terminal uridylyl transferase modifies the 3'-end of U6 small nuclear RNA. Nucleic Acids Research (1998) vol. 26 (13) pp. 3119-26

Trippe et al. Biochemical characterization of a U6 small nuclear RNA-specific terminal uridylyltransferase. Eur J Biochem (2003) vol. 270 (5) pp. 971-80

Trippe et al. Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. RNA (New York, NY) (2006) vol. 12 (8) pp. 1494-504

Tuschl et al. Targeted mRNA degradation by double-stranded RNA in vitro. Genes Dev (1999) vol. 13 (24) pp. 3191-7

Vadla, B., et al., lin-28 controls the succession of cell fate choices via two distinct activities. PLoS Genet, 2012. 8(3): p. e1002588.

Van Wynsberghe, P.M., et al., LIN-28 co-transcriptionally binds primary let-7 to regulate miRNA maturation in Caenorhabditis elegans. Nat Struct Mol Biol, 2011. 18(3): p. 302-8.

Viswanathan, S.R., et al., Lin28 promotes transformation and is associated with advanced human malignancies. Nat Genet, 2009. 41(7): p. 843-8.

Viswanathan, S.R., G.Q. Daley, and R.I. Gregory, Selective blockade of microRNA processing by Lin28. Science, 2008. 320(5872): p. 97-100.

Wang et al. DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet (2007) vol. 39 (3) pp. 380-5

Wang et al. Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. Nature (2008) vol. 456 (7224) pp. 921-6

Wang et al. Structure of the guide-strand-containing argonaute silencing complex. Nature (2008(a)) vol. 456 (7219) pp. 209-13

Watson and Crick. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature (1953) vol. 171 (4356) pp. 737-8

West, J.A., et al., A role for Lin28 in primordial germ-cell development and germ-cell malignancy. Nature, 2009. 460(7257): p. 909-13.

Wightman, B., I. Ha, and G. Ruvkun, Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell, 1993. 75(5): p. 855-62.

Wilbert et al. LIN28 binds messenger RNAs at GGAGA motifs and regulates splicing factor abundance. Molecular Cell, 2012. 48(2): p. 195-206.

Winter, J., et al., Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat Cell Biol, 2009. 11(3): p. 228-34.

Wu, L. and J.G. Belasco, Micro-RNA regulation of the mammalian lin-28 gene during neuronal differentiation of embryonal carcinoma cells. Mol Cell Biol, 2005. 25(21): p. 9198-208.

Wyman et al. Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. Genome Res (2011) pp.

Xu, B., K. Zhang, and Y. Huang, Lin28 modulates cell growth and associates with a subset of cell cycle regulator mRNAs in mouse embryonic stem cells. RNA, 2009. 15(3): p. 357-61.

Yang et al. Characterization of 3'hExo, a 3' exonuclease specifically interacting with the 3' end of histone mRNA. J Biol Chem (2006) vol. 281 (41) pp. 30447-54

Yates et al. Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase. Nature Structural & Molecular Biology (2012) vol. 19 (8) pp. 782-7

Yu, J., et al., Induced pluripotent stem cell lines derived from human somatic cells. Science, 2007. 318(5858): p. 1917-20.

Yuan, J., et al., Lin28b reprograms adult bone marrow hematopoietic progenitors to mediate fetal-like lymphopoiesis. Science, 2012. 335(6073): p. 1195-200.

Yue et al. CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: characterization of the CCA-adding enzyme from the archaeal hyperthermophile Sulfolobus shibatae. RNA (1996) vol. 2 (9) pp. 895-908

Zamore et al. RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. Cell (2000) vol. 101 (1) pp. 25-33

Zeng and Cullen. Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. J Biol Chem (2005) vol. 280 (30) pp. 27595-603

Zeng et al. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. EMBO J (2005) vol. 24 (1) pp. 138-48

Zhang et al. Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP. EMBO J (2002) vol. 21 (21) pp. 5875-85

Zhang et al. Single processing center models for human Dicer and bacterial RNase III. Cell (2004) vol. 118 (1) pp. 57-68

Zhang, W.C., et al., Glycine decarboxylase activity drives non-small cell lung cancer tumor-initiating cells and tumorigenesis. Cell, 2012. 148(1-2): p. 259-72.

Zhu, H., et al., Lin28a transgenic mice manifest size and puberty phenotypes identified in human genetic association studies. Nat Genet, 2010. 42(7): p. 626-30.

Zhu, H., et al., The Lin28/let-7 axis regulates glucose metabolism. Cell, 2011(a). 147(1): p. 81-94.

CHAPTER 2

Lin28-mediated control of let-7 expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7)

Text and figures partially adapted from "Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7)" by J.E. Thornton, E. Piskounova, H.M. Chang, and R.I. Gregory published in the RNA Journal (2012). The author is responsible for all data shown in Figures 2.1 and 2.2, and assisted in data shown in other figures. E. Piskounova performed experiments in Figure 2.3 and H.M. Chang performed experiments in Figure 2.4. All authors were instrumental in designing experiments and in providing helpful discussion. The author and R.I. Gregory wrote the manuscript.

Thornton et al. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). RNA (New York, NY) (2012) vol. 18 (10) pp. 1875-85

INTRODUCTION

microRNAs (miRNAs) are small, 22-nucleotide (nt) noncoding RNAs that repress the expression of many target messenger RNAs (mRNAs) (Wightman et al., 1993). The canonical process of miRNA biogenesis is well understood and is characterized by successive cleavage events by RNase III enzymes (Winter et al., 2009). After transcription by RNA polymerase II, primary miRNA transcripts (pri-miRNAs) are cleaved by the Microprocessor complex consisting of the RNAse III enzyme Drosha and its essential double-stranded RNA-binding partner DGCR8, yielding a short hairpin miRNA precursor (pre-miRNA) (Denli et al., 2004; Gregory et al., 2004; Han et al., 2006). Pre-miRNAs are exported to the cytoplasm by Exportin 5 where they are processed by Dicer, another RNase III enzyme, yielding the canonical 22nt miRNA duplex (mature miRNA) (Hutvagner et al., 2001). The mature miRNA undergoes strand selection where one strand of the duplex (guide strand) is preferentially incorporated into the miRNA-Induced Silencing Complex (miRISC) over the other strand (passenger or miR* strand). At its center, miRISC contains an Argonaute protein, and through base pairing between the mature miRNA and the 3'UTR of a target mRNA mediates translational inhibition and/or mRNA decay (Bartel, 2009; Fabian et al., 2010).

Proper temporal and spatial expression of miRNAs is essential for normal development and physiology, as perturbations in specific miRNAs or miRNA processing factors can lead to aberrant development and cancer (Calin and Croce, 2006; Esquela-Kerscher and Slack, 2006; Small and Olson, 2011). In embryonic cells, the RNA-binding protein Lin28 coordinately represses the let-7 family of miRNAs by binding to the terminal loop (also known as pre-element or preE) of pre- and pri-let-7 miRNAs, thereby inhibiting let-7 biogenesis (Heo et al., 2008; Newman et al., 2008; Rybak et al., 2008; Viswanathan et al., 2008). As cells undergo

differentiation Lin28 levels decrease, leading to a corresponding increase in mature let-7, which is retained in many adult tissues (Martinez and Gregory, 2010). Furthermore, Lin28 mRNA is repressed by let-7 miRNAs, leading to an inversely correlated expression pattern between let-7 and Lin28 and a double negative feedback loop that controls cell differentiation (Wu & Belasco, 2005). Lin28 is required for normal development and contributes to the pluripotent state by preventing let-7 mediated differentiation of embryonic stem cells (ESCs) (Ambros & Horvitz, 1984; Moss et al., 1997; Viswanathan & Daley, 2010). Lin28 overexpression or let-7 inhibition with antisense RNAs promotes reprogramming of human and mouse fibroblasts to induced pluripotent stem cells (iPSCs) (Ambros and Horvitz, 1984; Melton et al., 2010; Yu et al., 2007). The Lin28/let-7 axis is also relevant to a wide variety of human cancers as well as the control of glucose homeostasis in mammals (Frost and Olson, 2011; Iliopoulos et al., 2009; Piskounova et al., 2011; Viswanathan et al., 2009; Zhu et al., 2011).

In mammals there are two Lin28 paralogs, Lin28A (Lin28) and Lin28B. Lin28A recognizes pre-let-7 in the cytoplasm and recruits the terminal uridyl transferase (TUTase) Zcchc11 (TUTase4/TUT4) to add an oligouridine tail to the 3' end of pre-let-7, blocking Dicer cleavage and leading to the degradation of the pre-miRNA (Hagan et al., 2009; Heo et al., 2008; Heo et al., 2009). Lin28B is predominantly localized to the nucleus and blocks pri-miRNA processing through a TUTase-independent mechanism (Piskounova et al., 2011). Most cells that express a Lin28 family member do so selectively such that either Lin28A or Lin28B is exclusively expressed. This selective expression and differential localization of Lin28 family members allows for the repression of let-7 by distinct mechanisms in different cell and tumor types. The recently-identified TUTase Zcchc11 may also regulate IL-6 levels by uridylating mature miR-26a, promote the cell-cycle-dependent degradation of a subset of histone mRNAs,

and is required for the growth of Lin28A-driven cancers *in vitro* and *in vivo* (Jones et al., 2009; Piskounova et al., 2011; Schmidt et al., 2011). Given its central role in processes ranging from the inflammatory response, to cell cycle regulation and Lin28-mediated repression of let-7, Zcchc11 is an important RNA-modifying enzyme that may have essential roles in diverse aspects of human biology. However, very little is known about the *cis*-acting elements of mammalian TUTases or how TUTases interact with their binding partners.

In this study we set out to examine the mechanism by which Zcchc11 represses pre-let-7 in a Lin28-dependent manner. Mutational analyses of Zcchc11 identified domains required for activity both in the absence and presence of Lin28, and using recombinant proteins we were able to show that Lin28 and Zcchc11 proteins are sufficient for uridylation of pre-let-7 *in vitro*. Furthermore, we found that the single C2H2-type zinc finger at the N-terminus of Zcchc11 mediates the functional interaction with Lin28. Comparing the domain architecture of Zcchc11 to other mammalian TUTases we identified Zcchc6, another TUTase with extensive homology to Zcchc11, which also mediates Lin28-dependent uridylation of pre-let-7 *in vitro*. Accordingly, we found Zcchc6 depletion in embryonic cells synergized with Zcchc11 knockdown to upregulate let-7 miRNAs, implying that these two TUTases work redundantly to repress let-7 expression. These findings provide insight into the mechanism of Lin28-mediated TUTase control of let-7 expression in development, stem cells, and cancer.

RESULTS

Domains of Zcchc11 required for Lin28-enhanced pre-miRNA uridylation

Previous work from our group and others identified Zcchc11 as a cytoplasmic Lin28-interacting TUTase in embryonic and cancer cells (Hagan et al., 2009; Heo et al., 2009;

Piskounova et al., 2011). Its depletion in Lin28-expressing cells leads to the specific upregulation of let-7 family members similar to the depletion of Lin28 and its expression is required for potent let-7 repression and rapid cell growth in Lin28A-expressing cancers (Piskounova et al., 2011). Zcchc11 encodes a 184kDa non-canonical poly(A) polymerase that is highly conserved across vertebrates. The Zcchc11 active site is located within the Nucleotidyl Transferase (Ntr) domain, which is paired with a Poly(A)-Polymerase-Associated (PAP) domain, a common feature of non-canonical poly(A) polymerases (Kwak & Wickens, 2007; Martin & Keller, 2007; Saitoh et al., 2002). Catalysis requires a conserved aspartate triad in the Ntr and when a mutant lacking these residues is overexpressed, its functions is dominant negative (Hagan et al., 2009). Flanking the active site are three CCHC retroviral-type zinc fingers/zinc knuckles, which are implicated in nucleic acid binding. At the N-terminus of the protein is a region that shares significant homology with the active site, including a proximal PAP domain; however, this region lacks one of the crucial aspartates predicted to be necessary for catalysis. Instead, this region is most similar to the yeast TRF4 proteins, which carry out cytoplasmic poly(A) RNA polymerase activity (Saitoh et al., 2002). N-terminal to this region is a classical C2H2 zinc finger with no known function but these motifs are known to bind DNA, RNA, or protein. Finally, at the N- and C-termini of Zcchc11 there are two domains of unknown function that are similar to pneumoviridae attachment proteins and the glutamine-rich neurodegenerative disease-associated protein atrophin-1, respectively (Figure 2.1A).

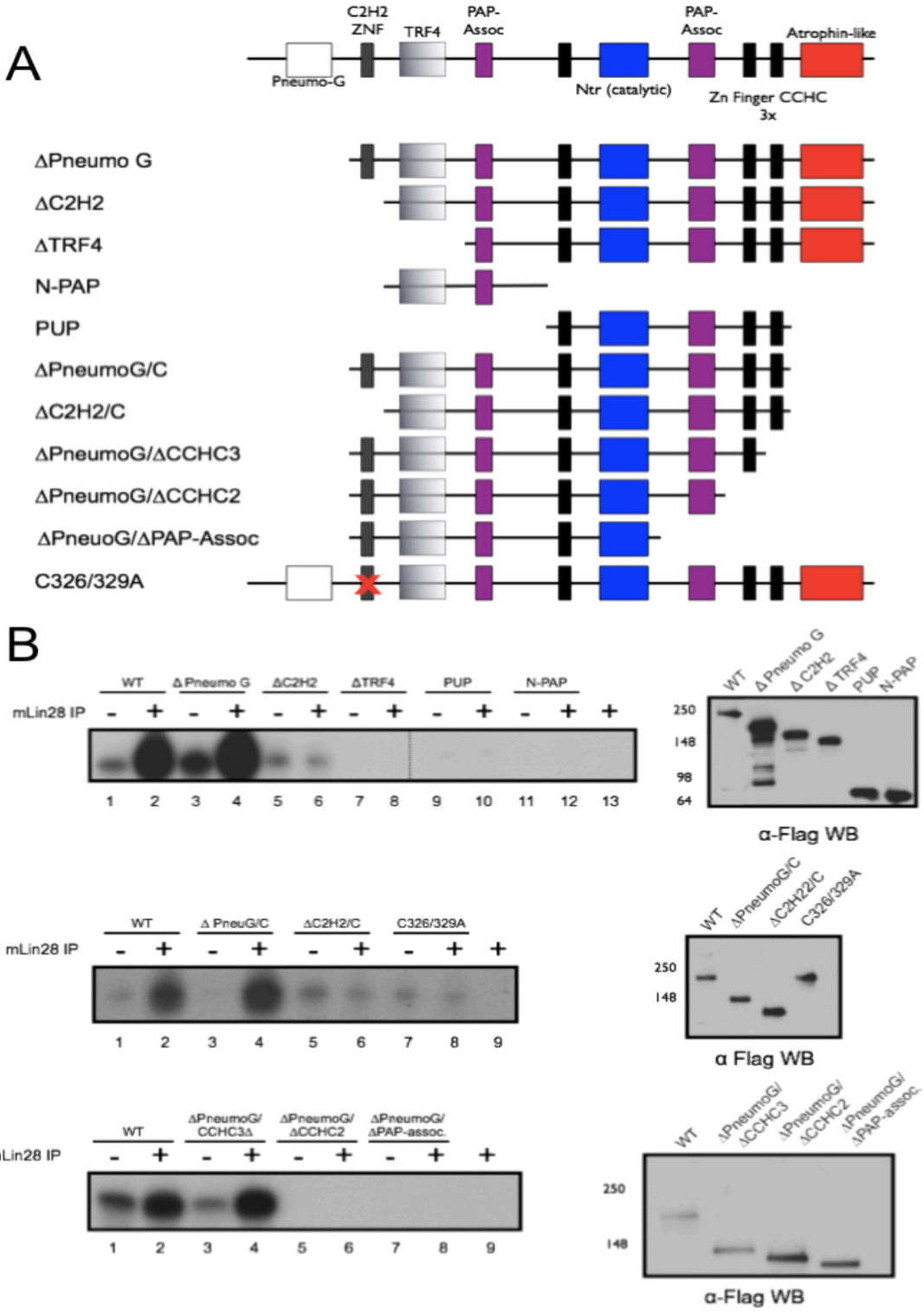
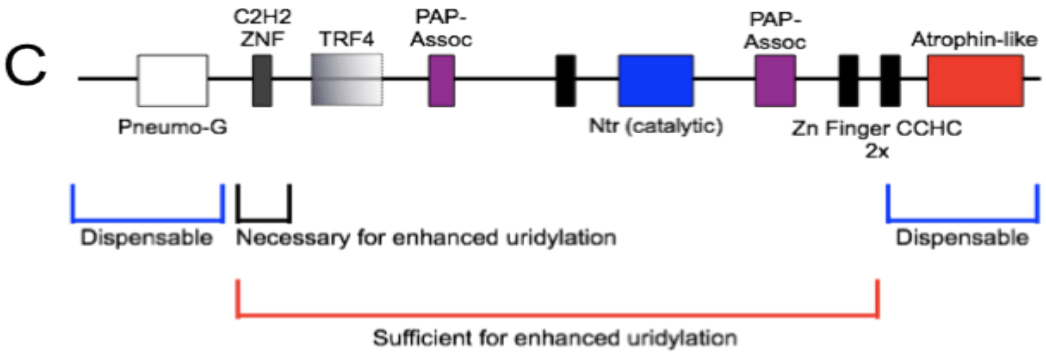
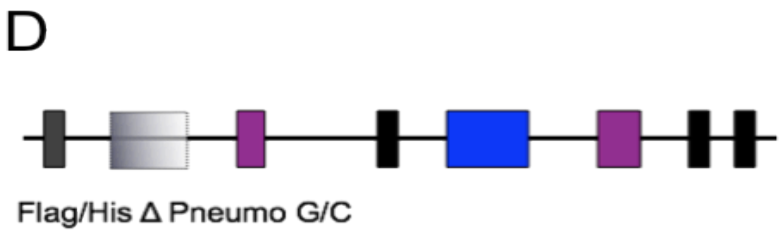


Figure 2.1: Domains of Zcchc11 required for Lin28-mediated pre-let-7 uridylation

Figure 2.1 continued





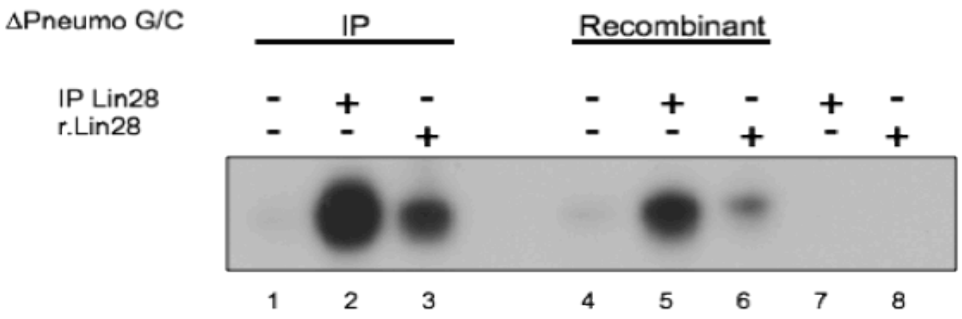


Figure 2.1 continued

A. Schematic representation of Zcchc11 and truncations used for *in vitro* uridylation assays.

B. Uridylation assays with synthetic pre-let-7g carried out using Flag immunopurified (IP) Zcchc11 variants and IP Lin28. a-Flag Western blots show similar amounts of IP Zcchc11 within experiments.

C. Summary of Zcchc11 domain requirements from *in vitro* uridylation assays.

D. Reconstitution of Lin28-enhanced pre-let-7 uridylation with recombinant proteins. Zcchc11 truncation ΔPneumoG/C purified from either HEK293T (IP) or E. coli (Recombinant) was incubated with either Flag-Lin28 (IP) or 6x-His Lin28 (r.Lin28) in a uridylation assay with synthetic pre-let-7g. (Top) Schematic representation of the domains present in ΔPneumoG/C.

To understand which domains of Zcchc11 are required for uridylation activity, we generated a series of mutant cDNAs and tested the ability of the resulting Flag-immunopurified (Flag IP) proteins to uridylate synthetic pre-let-7 miRNA *in vitro*. Mutants were generated lacking N- and C-terminal domains or harboring point mutations in conserved residues (Figure 2.1A). As described previously, Zcchc11 exhibits a low-level of uridylation against pre-let-7 and this activity is strongly enhanced by the addition of recombinant or immunopurified Lin28 (Hagan et al., 2009; Heo et al., 2009). As determined by the incorporation of radiolabeled UTP, the activity of wild type Zcchc11 was compared for each of the mutants (Figure 2.1B).

We found that after performing serial N-terminal truncations, the pneumoviridae (PneumoG) domain was dispensable for both basal level uridylation and activity enhanced by IP Flag-Lin28. Loss of the N-terminal C2H2 domain still allowed for basal activity, but this mutant could no longer support Lin28-enhanced uridylation against pre-let-7, indicating that the C2H2 zinc finger may be essential for the interaction between Lin28 and Zcchc11. Interestingly, when the TRF4 domain was deleted we could not detect either basal or Lin28-enhanced activity. This result was surprising given the prediction that the TRF4 domain was insufficient to carry out catalytic activity on its own. Indeed, when we tested a fragment of Zcchc11 containing the N-terminal TRF4 domain but lacking the NTR domain, no uridylation activity was detected (Figure 2.1B, upper panel, compare lanes 7 and 8 to 11 and 12). All further mutant proteins tested lacking TRF4 failed to support any detectable uridylation activity (Figure 2.1B, upper panel, lanes 9 and 10 and data not shown).

To determine if our findings on N-terminal deletions of Zcchc11 could be supported in the context of additional C-terminal truncations, we tested mutants lacking the C-terminal Atrophin-like domain in combination with ΔPneumoG and ΔC2H2 mutants. The Atrophin-like

domain was dispensable in these experiments, indicating that it is not required for basal or Lin28-enhanced uridylation by Zcchc11. To confirm that the C2H2 zinc finger *per se* was required for Lin28-enhanced uridylation, a full-length Zcchc11 cDNA was generated bearing cysteine to alanine mutations in the residues predicted to be central to the C2H2 zinc finger (C326/329A). Indeed, this mutant exhibited only basal uridylation activity, as the addition of Lin28 had no impact on its catalysis *in vitro*. Given that this mutant phenocopied the ΔC2H2 and ΔC2H2/C mutants, we conclude that this zinc finger is required for Lin28-enhanced uridylation *in vitro*.

To define the minimal Zcchc11 mutant that supports Lin28-enhanced uridylation, we further examined the requirements of C-terminal domains. Compared to WT, a mutant lacking the C-terminal-most CCHC zinc finger exhibited robust basal and Lin28-enhanced activity, whereas additionally truncating the adjacent CCHC zinc finger led to no detectable activity, implying that the three CCHC zinc fingers may be required for different aspects of RNA recognition or positioning (Figure 2.1B, lower panel). These studies provide insight into the basic mechanism underlying the catalytic nature of Zcchc11 (Figure 2.1C).

In vitro reconstitution of Lin28-mediated pre-let-7 uridylation with recombinant proteins

The experiments described above suggest that specific domains of Zcchc11 mediate the interaction with Lin28 to uridylate pre-let-7 *in vitro*. To confirm that these two proteins are sufficient for activity and do not rely on contaminating or accessory factors interacting with the immunopurified proteins, we purified 6x-His Lin28 (r.Lin28) and Flag/6x-His ΔPneumoG/C Zcchc11 from E. coli. Compared to immunopurified ΔPneumoG/C Zcchc11, the Flag/6x-His protein uridylated pre-let-7 at the basal level to a similar extent indicating that the Zcchc11

expressed and purified from bacteria is catalytically active (Figure 2.1D). Adding either immunopurified Flag-Lin28 or recombinant 6x-His Lin28 to the reaction similarly enhanced the uridylation of pre-let-7 by either Zcchc11 preparation. Neither of the Lin28 proteins themselves led to detectable levels of uridylated pre-let-7, indicating that labeled products originated from the enzymatic activity of Zcchc11. These experiments show that the combination of Lin28 and Zcchc11 proteins are necessary and sufficient to carry out the robust uridylation of pre-let-7 *in vitro* (Figure 2.1D).

The let-7 preE confers Lin28-enhanced pre-miRNA uridylation by Zcchc11

To understand the role of pre-miRNA substrates in Zcchc11-mediated uridylation, we investigated which *cis*-acting RNA elements support uridylation enhanced by Lin28. Lin28 binding to pre-let-7 requires specific sequence and structural information in both the RNA and the protein. The cold-shock domain (CSD) of Lin28 is inserted into the terminal loop of various pre-let-7 RNAs, and the Lin28 CCHC zinc fingers dimerize to recognize a GGAG motif proximal to the Dicer cleavage site of pre-let-7 (Loughlin et al., 2011; Nam et al., 2011). The various domains of pre-let-7 required for Zcchc11-mediated uridylation are, however, to this point, unknown.

To determine the regions of pre-let-7 required for uridylation by Zcchc11, we took advantage of the understanding that the let-7 preE is bound by recombinant Lin28 as efficiently as full-length pre-let-7 (Piskounova et al., 2008). If Lin28 binding is sufficient to direct Zcchc11-mediated uridylation, then RNA substrates with divergent sequences outside of the let-7 preE should be comparable substrates to pre-let-7. To test this, we generated two synthetic pre-

miRNAs; one composed of the preE of let-7g and the stem sequence of miR-21 (pre-21S7L), and another composed of the preE of miR-21 and the stem of let-7g (pre-7S21L) (Figure 2.2A).

We compared the uridylation activity of Zcchc11 towards these chimeric RNAs versus both pre-let-7g and pre-miR-21. As shown in Figure 2.2B, pre-let-7g undergoes robust uridylation with the addition of IP Flag-Lin28. miR-21 is uridylated at a basal level similar to that of pre-let-7, but the addition of IP Flag-Lin28 has no effect on uridylation levels as described previously (Hagan et al., 2009). Zcchc11 also only exhibits basal activity towards pre-7S21L, however, when the chimeric pre-21S7L is incubated with IP Flag-Lin28, it is subjected to enhanced uridylation activity similar to that of WT pre-let-7 (Figure 2.2B, compare lanes 7 and 8 to 1 and 2). This result suggests that the effect of Lin28 binding to the preE of pre-let-7 is sufficient to allow targeting and uridylation by Zcchc11.

To further examine the RNA determinants supporting Zcchc11 and Lin28-mediated uridylation, we monitored the substrate preference of several other synthetic RNAs in *in vitro* uridylation assays. One explanation for the sufficiency of the let-7 preE to support uridylation is that pre-miRNA stem regions are dispensable altogether and the loop alone encodes all necessary regulatory information. To test this hypothesis we performed uridylation assays on pre-let-7g, the preE of let-7g alone, or a chimeric RNA bearing the preE of let-7g with only the 3' stem region of miR-21 (pre-S21L7Δ5). Compared to pre-let-7g, neither let-7g preE nor pre-S21L7Δ5 underwent Lin28-enhanced uridylation, indicating the necessity of an intact pre-miRNA stem to drive this activity (Figure 2.2C).

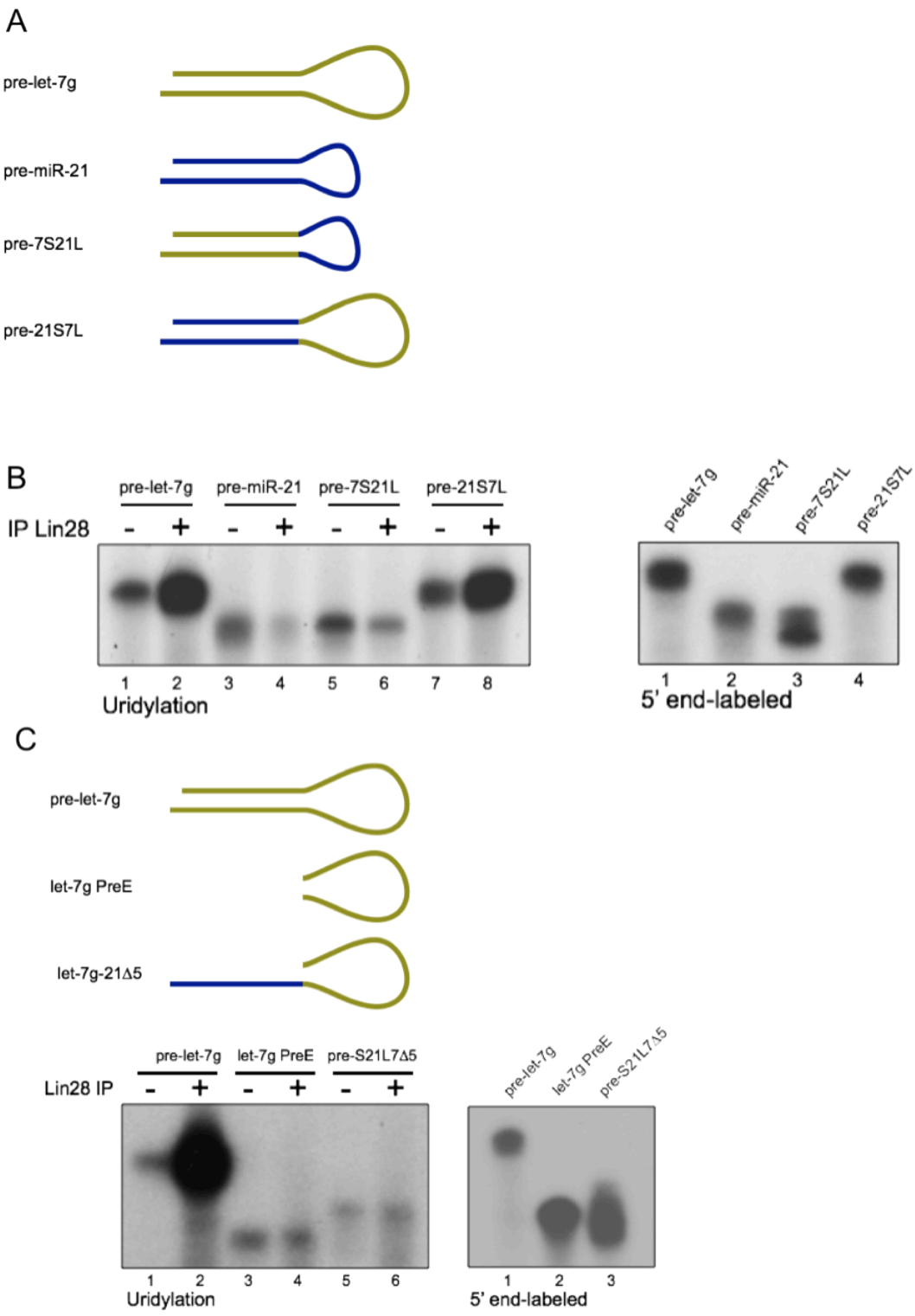


Figure 2.2: The preE of let-7 is sufficient to direct both Lin28 binding and uridylation of pre- let-7

Figure 2.2 continued

A. Diagram of synthetic RNAs used for *in vitro* uridylation assays. (Pre-let-7g) Endogenous precursor let-7g miRNA sequence; (Pre-miR-21) endogenous precursor miR-21 sequence; (Pre-7S21L) synthetic RNA consisting of the miR-21 preE and let-7g stem; (Pre- 21S7L) synthetic RNA consisting of the let-7g preE and miR-21 stem sequences.

B. (left) Uridylation assay as in Figure 1 using WT Flag IP-mZcchc11, with or without Flag IP-mLin28, and the indicated precursor miRNAs. (Right) 59 end-labeled RNAs showing equal amounts.

C. (Top) Diagram of synthetic RNAs used for *in vitro* uridylation assays. (Bottom left) Uridylation assays and (bottom right) 59 end-labeled RNAs showing equal amounts as in (B).

The related TUTase Zcchc6 is functionally redundant with Zcchc11 *in vitro*

Our findings on the domains of Zcchc11 supporting Lin28-mediated uridylation *in vitro* led us to examine other TUTases as potential regulators of pre-miRNAs. Among the seven non-canonical poly(A) polymerases encoded in the human genome, we found that Zcchc6 (PAPD6/TUTase 7) has striking homology to Zcchc11 including the domains constituting its active site, its three CCHC zinc fingers, the N-terminal TRF4/PAP-associated domains, and C2H2 zinc finger (Figure 2.3A). Importantly, there is extensive conservation between Zcchc11 and Zcchc6 at critical residues in the active site and in the C2H2 zinc finger (Figure 2.3A-C). To determine if Zcchc6 shares activity similar to Zcchc11, we tested the ability of IP Flag-hZcchc6 to uridylate pre-let-7 *in vitro* in the absence or presence of Lin28. Similar amounts of Flag-hZcchc11 or Flag-hZcchc6 were used in uridylation assays with Flag-hLin28A and both TUTases were stimulated to an equal extent (Figure 2.3D, E). To confirm that the effects seen with Flag-hLin28A were not dependent on the paralog of Lin28 used, we also tested the stimulatory effect of Flag-hLin28B since both Lin28 proteins act identically *in vitro* (Figure 2.3D-F) (Heo et al., 2009).

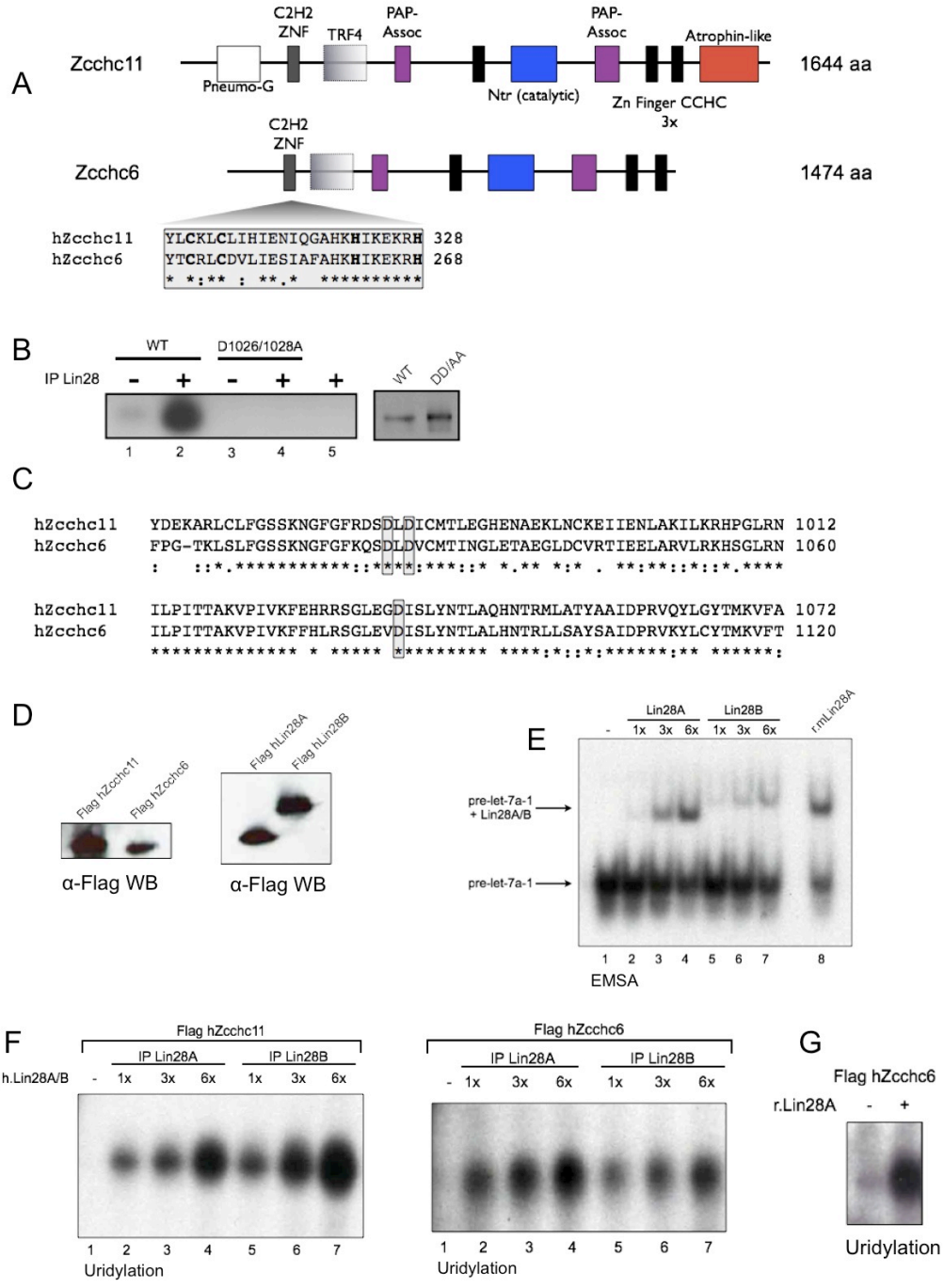


Figure 2.3: Zcchc11 and Zcchc6 have a highly similar domain organization and can both mediate Lin28-dependent pre-let-7 uridylation *in vitro*

Figure 2.3 continued

A. Schematic showing the domain similarities between hZcchc11 and hZcchc6 with the N-terminal C2H2 zinc finger highlighted and critical zinc finger residues in bold.

B. Uridylation assay with Flag-IP WT mZcchc11 and a mutant harboring point mutations in two conserved aspartates required for catalysis, with or without Flag-IP Lin28.

C. Alignment of the nucleotidyl transferase (Ntr) domains of hZcchc11 and hZcchc6. Aspartic acid residues critical for catalysis are boxed.

D. α-Flag WB showing relative amounts of Flag-hZcchc11 and Flag-hZcchc6 (left) or Flag-hLin28A and Flag-hLin28B (right).

E. EMSA showing similar amounts of functional Flag-hLin28A and Flag-hLin28B used in uridylation assays.

F. Uridylation assays using Flag-hZcchc11 or Flag-hZcchc6 with either Flag-hLin28A or Flag-hLin28B.

G. Uridylation assay with Flag-hZcchc6 and r.Lin28A.

In these experiments either Lin28A or Lin28B enhanced the uridylation activity of either TUTase in a dose-dependent manner. The enhancement in hZcchc6 uridylation activity was also observed using r.Lin28, indicating that the this effect was not due to co-immunoprecipating proteins and that Zcchc6 and Zchc11 are functionally indistinguishable in these assays (Figure 2.3G).

These results *in vitro* suggest both TUTases may recognize let-7 precursors in biologically relevant settings. Zcchc6 has previously been shown to have poly(U) activity *in vitro* (Kwak & Wickens, 2007; Rissland et al., 2007), and depletion of Zcchc6 in colon cancer cells led to reduced levels of uridylated mature let-7e (Wyman et al., 2011). Zcchc6 is also a homolog of *C. elegans* CDE-1, which uridylates a subset of siRNAs bound by the Argonaute protein CSR-1 and loss of CDE-1 leads to aberrant chromosomal segregation and dysregulation of CSR-1-bound siRNAs (van Wolfswinkel et al., 2009). In spite of these data, this is the first evidence of Zcchc6 uridylating pre-miRNAs and suggests parallel activity with Zcchc11 and a role in the Lin28 pathway.

Zcchc11 and Zcchc6 redundantly control let-7 biogenesis in embryonic stem cells

Given the findings on the activity of Zcchc6 *in vitro*, we investigated whether Zcchc6 functions in parallel with Zcchc11 *in vivo*. We have previously shown that Zcchc11 depletion in embryonic carcinoma (EC) and embryonic stem (ES) cells led to the coordinate derepression of let-7 miRNAs, while Zcchc6 depletion led to no change in mature let-7 levels (Hagan et al., 2009; Heo et al., 2009). The derepression observed upon Zcchc11 knockdown was, however, generally more modest than the depletion of Lin28 in all cell types tested (Hagan et al., 2009; Heo et al., 2009; Piskounova et al., 2011). One interpretation of this finding is that there are

redundant factors working in parallel with Lin28 and Zcchc11 to repress let-7 miRNAs in undifferentiated cell types. To test whether Zcchc6 works redundantly with Zcchc11, we used siRNAs to deplete both TUTases in P19 and V6.5 cell lines (EC and ES cells, respectively). Upon Zcchc11 knockdown with two independent siRNAs there was a modest 2- to 3-fold upregulation of mature let-7g, as we have previously shown, whereas consistent with previous reports, depletion of Zcchc6 with two independent siRNAs led to no significant changes at the level of mature let-7g. When both TUTases were knocked down, however, we observed a consistent upregulation in mature let-7 that was more dramatic than either individual knockdown alone (Figure 2.4). This trend was specific to let-7 family members, as levels of the unrelated miRNA miR-21 were unchanged. Moreover global profiling revealed that changes in miRNA expression were restricted to let-7 family members (data not shown). This trend was seen in both P19s and V6.5s, suggesting Zcchc11/Zcchc6 redundancy is a general mechanism of embryonic cells. The synergistic relationship between these two related TUTases, both *in vitro* and *in vivo*, could explain the modest effects seen for depletion of either TUTase individually in Lin28-expressing cells and expands the repertoire of miRNA modifying enzymes.

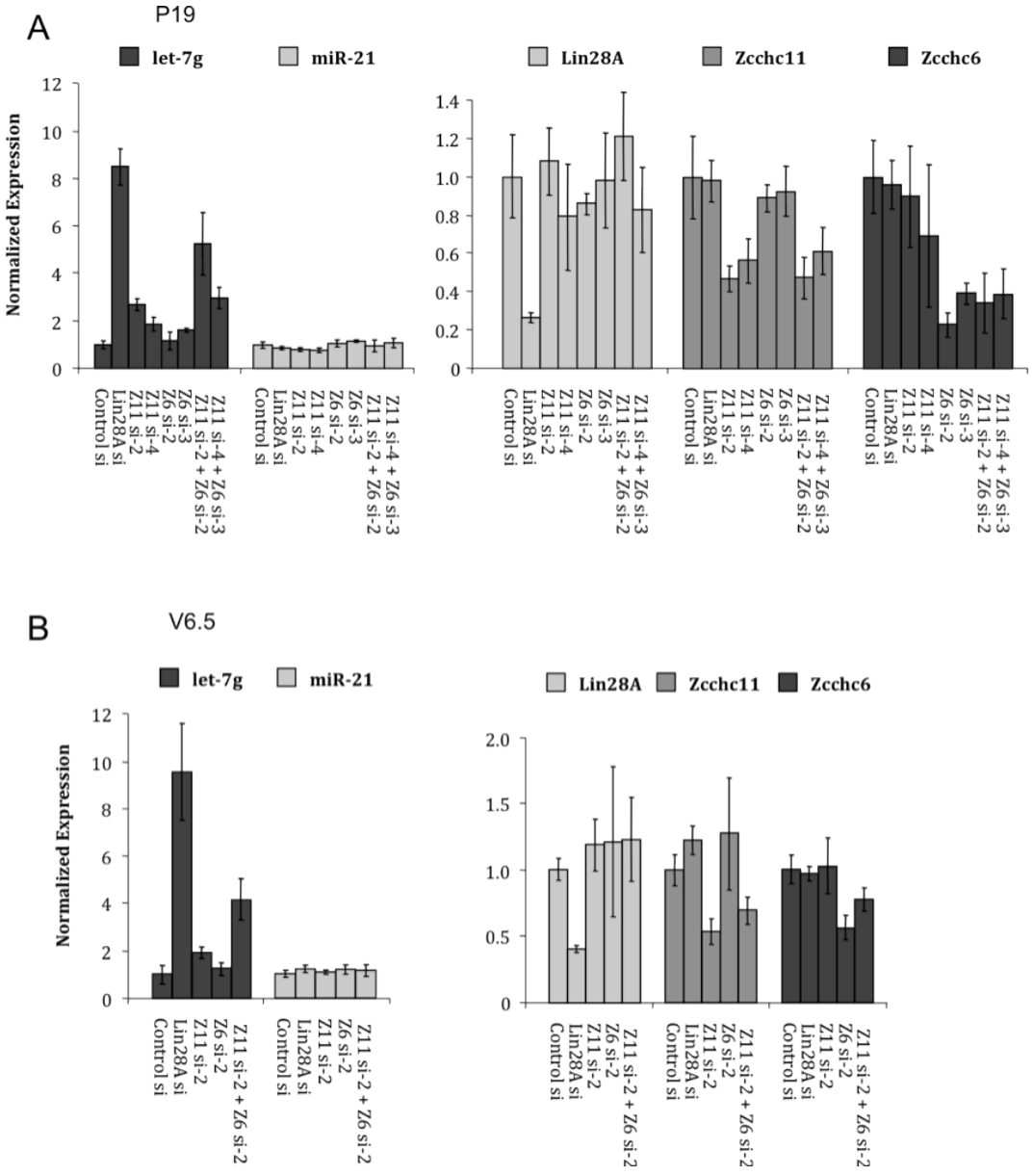


Figure 2.4: Zcchc11 and Zcchc6 function redundantly to suppress let-7 expression in embryonic cells

Figure 2.4 continued

A. qRT–PCR analysis of mature let-7g and mature miR-21 levels in P19 embryonal carcinoma cells after transfection with the indicated siRNAs (left). mRNA levels of the indicated genes in P19 cells after transfection with the indicated siRNAs (right).

B. qRT–PCR as in A in V6.5 mouse embryonic stem cells. For all experiments, miRNA levels were normalized to sno-142 and mRNA levels were normalized to b-actin. Error bars represent SD of experiments in triplicate.

DISCUSSION

Recent work examining the role of miRNAs in development and cancer has revealed extensive post-transcriptional control at various levels of miRNA biogenesis (Siomi & Siomi, 2010). Lin28 (Lin28A) and Lin28B have emerged as important posttranscriptional regulators of let-7 expression in stem cells, development, metabolism, and disease (Viswanathan & Daley, 2010). In the case of Lin28A, this regulation involves the recruitment of a TUTase Zcchc11 to catalyze the 3' terminal uridylation of pre-let-7 RNAs. Several studies have identified extensive non-templated nucleotide addition to the 3' ends of mature and precursor miRNAs (Ameres et al., 2010; Berezikov et al., 2011; Burroughs et al., 2010; Chiang et al., 2010; Heo et al., 2008; Jones et al., 2009; Katoh et al., 2009; Lehrbach et al., 2009; Newman et al., 2011). Our work here is the first extensive mechanistic analysis of one of these enzymes, Zcchc11. First, we have uncovered specific domains that are required for mediating efficient Lin28-endhanced uridylation of pre-let-7 *in vitro*. Of the four zinc fingers encoded in Zcchc11, the unique C2H2 zinc finger at the N-terminus of the protein mediates the functional interaction with Lin28, as point mutations in conserved cysteine residues of this zinc finger abolish Lin28-enhanced uridylation activity. The TRF4 domain at the N-terminus of Zcchc11, while incapable of supporting uridylation activity on its own, is nonetheless required for activity *in vitro*. This essential role may explain its significant degree of conservation across taxa. Furthermore, the CCHC zinc fingers, which define a class of at least 13 mammalian proteins, are differentially required for uridylation activity *in vitro*. Specifically, the C-terminal-most CCHC zinc finger is dispensable for *in vitro* activity, while the zinc finger just C-terminal to the active site is required for any detectable activity. Finally, there are regions dispensable for Lin28-enhanced uridylation at the N- and C-termini of Zcchc11 (Figure 2.1C). Both of these domains are of unknown

function but remain conserved in other organisms. Since Zcchc11 has been implicated in several other biological pathways, we cannot rule out that the domains identified as dispensable are required for other processes. Indeed, a recent study has identified the N-terminal portion of Zcchc11, which lacked any catalytic regions to be sufficient to alter the cell cycle of cultured human cancer cells (Blahna et al., 2011). Our study unveils critical domains and residues that are required for Lin28-dependent Zcchc11 activity. Though still controversial, the Lin28-mediated control of let-7 expression in *C. elegans* has also been reported to involve pre-let-7 uridylation (Lehrbach et al., 2009; Van Wynsberghe et al., 2011). Notably however, the proposed Zcchc11 ortholog, PUP-2, lacks the C2H2 domain that we find mediates the functional interaction between Lin28 and Zcchc11 (Lehrbach et al., 2009). Therefore it remains unclear if and how Lin28 in worms recruits PUP-2 to repress let-7 expression.

We also examined the requirements of pre-let-7 that mediate Lin28-enhanced activity and we found the preE of let-7g in the context of an intact pre-miRNA to be sufficient to direct this activity. Although Zcchc11 recognizes and uridylates the 3' end of pre-let-7 family members and other miRNAs, this occurs through a mechanism that is independent of sequence information proximal to the site of uridylation. Instead, Lin28 bound to an intact preE sequence is sufficient to direct robust uridylation of the pre-miRNA. While the preE in our studies contains the Lin28-binding motif of GGAG, this sequence was previously shown to be insufficient in directing uridylation activity towards pre-let-7, as gain-of-function experiments indicated the positioning of the motif relative to the Dicer cleavage site was also essential (Heo et al., 2009). The chimeric pre-21S7L, however, has the GGAG motif positioned not in the preferred position ending 4nt before the Dicer cleavage site, but only 2nt away from this point, suggesting there are other sequence or structural determinants directing Zcchc11-mediated uridylation against pre-let-7

miRNAs. Furthermore, although the preE of let-7 is sufficient for Lin28-enhanced uridylation activity, an intact pre-miRNA structure is required, since a viable substrate lacking one arm of the pre-miRNA stem was no longer targeted for robust uridylation. This requirement likely comes from RNA-protein interactions between Zcchc11 and duplex-form RNA because recent structural studies have shown the conserved domains of Lin28 interacting exclusively with the PreE of several let-7 family members. The roles of the protein domains of Zcchc11 mediating this dsRNA interaction warrant further research. We rule out the role of other protein factors giving further specificity to pre-let-7 uridylation since the reaction could be reconstituted from recombinant proteins produced in bacteria, but what defines this level of specificity remains unknown. Recent structural studies have uncovered the degree to which the let-7 preE is altered by Lin28 binding, revealing a partial unwinding of the duplex region near the site of Dicer cleavage (Nam et al., 2011). While it is unknown how far this melting proceeds into the stem of pre-let-7, this structural change could alter the RNA so that it is a preferred substrate of Zcchc11. Indeed, a recent structural study showed that the CCHC zinc fingers of Lin28 preferentially bind the single stranded heptad sequence of AGGAGAU in the stem of pre-let-7, providing evidence of sequence-specific RNA binding by zinc finger-containing proteins (Loughlin et al., 2011). Alternatively, pre-let-7-bound Lin28 may undergo a conformational change and this may provide a suitable protein-protein interaction surface between the Lin28-let-7 complex and Zcchc11 (Nam et al., 2011). More detailed RNA mutagenesis and/or structural studies examining the interplay between pre-let-7 and these two RNA binding proteins should provide additional insight into precisely how Lin28 functionally enhances Zcchc11 recognition of the let-7 preE. The uridylation and adenylation of mature miRNAs by Zcchc11 has also been reported (Jones et al., 2009; Wyman et al., 2011). Since Zcchc11 exhibits similar basal activity towards unrelated

pre-miRNAs (Figure 2.2) there may be other sequence-specific recognition factors that guide Zcchc11 activity towards other RNA substrates including mature miRNAs.

The findings in our mutational analysis led us to investigate other putative TUTases, and to the identification of Zcchc6 as a regulator of let-7 expression. One study investigating the potential redundancy between Zcchc11 and Zcchc6 found that only Zcchc11 was capable of binding stem-loop containing histone mRNAs, while Zcchc6 appeared to lack this capacity (Schmidt et al., 2011). In the case of Lin28 and let-7, however, we found Zcchc6 functioned identically to Zcchc11 *in vitro* as its enzymatic activity against a synthetic let-7 precursor was enhanced by either Lin28A or Lin28B, as was previously shown for Zcchc11 (Heo et al., 2009). We furthermore found Zcchc6 to be crucial in efficiently repressing mature let-7 miRNAs in embryonic cells. Although the double knockdown of Zcchc11 and Zcchc6 led to more dramatic let-7 derepression than the loss of either TUTase alone, it still did not reach the levels observed upon Lin28A knockdown. This could be explained by incomplete knockdown of both TUTases or the activity of other as-yet unidentified let-7 repressive factors. The identification of a second TUTase regulating let-7 turnover may provide valuable insight into the control of let-7 expression in cancer and embryonic stem cell biology. The expression pattern and localization of Zcchc6 are unknown but we anticipate that the relative expression levels of these two redundant TUTases will determine the relative contribution of Zcchc6 and Zcchc11 in the Lin28A-mediated control of let-7 expression. In this regard, we have recently shown that Zcchc11 inhibition in Lin28A-driven cancers can block tumor growth *in vitro* and *in vivo* (Piskounova et al., 2011). It will be important to explore the relevance of Zcchc6 in this context. To our knowledge there have been no studies examining expression patterns of these two TUTases in human disease, and it will be important to determine if either or both of these enzymes is

correlated with malignancies characterized by repressed let-7 levels and/or increased Lin28A expression. Whereas Lin28 proteins are likely difficult chemotherapeutic targets due to their non-enzymatic activity, Zcchc11 (and presumably Zcchc6) poses an intriguing possibility as a drug target because of its defined active site and the available structural data regarding non-canonical poly(A) polymerases (Stagno et al., 2010). Moreover, our ability to reconstitute this regulatory pathway with recombinant proteins provides an opportunity to perform *in vitro* screening to identify small molecule inhibitors of TUTase activity as potential new chemotherapeutic agents. These possibilities pose future areas of study and expand upon the novel centrality of uridylation in stem cell maintenance and tumor development.

MATERIALS AND METHODS

Cloning

All mZcchc11 mammalian-expression mutants were cloned into the XhoI and SalI sites of pBK_2x Flag EF1 vector. C326/329A mutant Zcchc11 was generated by site-directed mutagenesis using the QuickChange kit (Stratagene). hZcchc6 was amplified from HEK293 cDNA and cloned into the HindIII and BamHI sites of pFLAG-CMV2 (Sigma). For recombinant protein expression ΔPneumoG/C mZcchc11 was cloned into the SalI and NotI sites of pETDUET-1. Expression constructs for Flag-m.Lin28A, Flag-Lin28A, Flag-Lin28B, recombinant His-Lin28A, Flag-hZcchc11 wild-type and D1026/1028A mutant were described previously (Hagan et al., 2009; Piskounova et al., 2011; Piskounova et al., 2008; Viswanathan et al., 2008).

Immunoprecipitation and Recombinant Protein Production

Expression plasmids for Flag-Zcchc11, Flag-Zccch6, Flag-Lin28A, or Flag-Lin28B were transfected into HEK293T cells using Lipofectamine 2000 (Invitrogen). Cells were harvested in lysis buffer (20 mM Tris-HCl, pH 8.0, 137 mM NaCl, 1 mM EDTA, 1% Triton x100, 10% Glycerol, 1.5 mM $MgCl_2$, 5 mM DTT, 0.2 mM PMSF). Protein was purified using anti-Flag M2 beads (Sigma), eluted using Flag peptide (Sigma) and confirmed by Western Blot analysis, with a mouse anti-Flag antibody (Sigma). For recombinant protein production: Transformed BL21-CodonPlus® Competent bacteria (Stratagene) were grown to an OD^{600nm} of 0.4-0.6. Recombinant protein expression (r.Lin28A, and r.Zcchc11) was induced with 100μM IPTG for 2-3 hours. Cell pellets were resuspended in cold lysis buffer [20mM imidazole pH 8.0 in PBS, 0.1% Phenylmethyl sulfonyl fluoride (PMSF)] and sonicated. Cleared lysates were incubated

with Ni-NTA beads and after 90 minutes incubation at 4°C the beads were washed with 80 bead volumes wash buffer [10mM Tris (pH 7.8), 50mM imidazole pH 8.0, 500mM NaCl, 0.1% PMSF, 1mM DTT). Bound His-tagged proteins were eluted from the column with 1 volume elution buffer [10mM Tris (pH 7.8), 500mM imidazole pH 8.0, 500mM NaCl, 1mM DTT, 0.1% fresh PMSF] and dialyzed overnight against BC100 [20 mM Tris-HCl (pH 7.8), 100 mM KCl, 0.2 mM EDTA, 10% glycerol]. Proteins were further purified by size exclusion chromatography using a Superose 6 gel filtration column (20 mM Tris-HCl (pH 7.8), 500 mM KCl, 0.2 mM EDTA, 0.2% NP40, 10% glycerol) and peak fractions were dialyzed overnight against BC100 and stored at 4°C.

***In vitro* uridylation assay**

Purified proteins were incubated with 4 pmol of unlabelled synthetic RNA (Dharmacon) for 1 hour at 37°C in a 30 μl reaction mixture containing 100 mM KCl, 20 mM Tris-HCl pH 7.6, 10% Glycerol, 125 nM [α-^{32}P]UTP, 3.2 mM $MgCl_2$, 40U RNasin ribonuclease inhibitor (Promega). Products were resolved on 15% denaturing polyacrylamide gels and bands were detected by autoradiography.

Electrophoretic mobility shift analysis

EMSA with purified His-Lin28A was performed with end-labeled synthetic pre-let-7 as described but without competitor yeast tRNA (Piskounova et al., 2008). Briefly, reactions were set up in binding buffer [50mM Tris, (pH7.5), 100mM NaCl, 10mM βMe, 20U RNasin (Promega)] with 0.5nM or 5 nM end-labeled pre-let-7g and incubated for 60 min at room

temperature. Bound complexes were resolved on native 5% polyacrylamide gels and visualized by autoradiography.

In vivo Knockdowns and quantitative RT-PCR

The indicated siRNAs (see supplemental Table 2**)** were reverse transfected in either P19 or feeder-free V6.5 mouse embryonic stem cells using Lipofectamine2000 in 6 well plates, according to the manufacturer's protocol (Invitrogen). Total RNA was isolated 60 hours post-transfection using TriZol reagent (Invitrogen). To analyze relative mRNA levels, 2μg of total RNA was reverse transcribed using random hexamers and SuperScriptIII (Invitrogen). miRNAs were reverse transcribed from 10ng total RNA using gene-specific stem-loop RT primers (Applied Biosystems). Relative levels of miRNAs were determined by TaqMan based real-time PCR, snoRNA-142 for normalization. For quantitative analysis of mRNA levels real-time RT-PCR was performed with either SYBR green or Taqman assays. Actin was used as control. For global microRNA profiling, the TaqMan Rodent MicroRNA A Array v2.0 was used with 350ng total RNA as starting material for the multiplex RT with pre-amplification, according to manufacturer's directions (Applied Biosystems). The resulting data was normalized to the U6 snRNA.

REFERENCES

Ambros, V., and Horvitz, H.R. (1984). Heterochronic mutants of the nematode Caenorhabditis elegans. Science *226*, 409-416.

Ameres, S.L., Horwich, M.D., Hung, J.H., Xu, J., Ghildiyal, M., Weng, Z., and Zamore, P.D. (2010). Target RNA-directed trimming and tailing of small silencing RNAs. Science *328*, 1534-1539.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Berezikov, E., Robine, N., Samsonova, A., Westholm, J.O., Naqvi, A., Hung, J.H., Okamura, K.,

Dai, Q., Bortolamiol-Becet, D., Martin, R.*, et al.* (2011). Deep annotation of Drosophila melanogaster microRNAs yields insights into their processing, modification, and emergence. Genome Res *21*, 203-215.

Blahna, M.T., Jones, M.R., Quinton, L.J., Matsuura, K.Y., and Mizgerd, J.P. (2011). Terminal uridyltransferase enzyme zcchc11 promotes cell proliferation independent of its uridyltransferase activity. J Biol Chem *286*, 42381-42389.

Burroughs, A.M., Ando, Y., de Hoon, M.J., Tomaru, Y., Nishibu, T., Ukekawa, R., Funakoshi, T., Kurokawa, T., Suzuki, H., Hayashizaki, Y.*, et al.* (2010). A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. Genome Res *20*, 1398-1410.

Calin, G.A., and Croce, C.M. (2006). MicroRNA signatures in human cancers. Nat Rev Cancer *6*, 857-866.

Chiang, H.R., Schoenfeld, L.W., Ruby, J.G., Auyeung, V.C., Spies, N., Baek, D., Johnston, W.K., Russ, C., Luo, S., Babiarz, J.E.*, et al.* (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev *24*, 992-1009.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231-235.

Esquela-Kerscher, A., and Slack, F.J. (2006). Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer *6*, 259-269.

Fabian, M.R., Sonenberg, N., and Filipowicz, W. (2010). Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem *79*, 351-379.

Frost, R.J., and Olson, E.N. (2011). Control of glucose homeostasis and insulin sensitivity by the Let-7 family of microRNAs. Proc Natl Acad Sci U S A.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235-240.

Hagan, J.P., Piskounova, E., and Gregory, R.I. (2009). Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nat Struct Mol Biol *16*, 1021-1025.

Han, J., Lee, Y., Yeom, K.H., Nam, J.W., Heo, I., Rhee, J.K., Sohn, S.Y., Cho, Y., Zhang, B.T., and Kim, V.N. (2006). Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell *125*, 887-901.

Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol Cell *32*, 276-284.

Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell *138*, 696-708.

Hutvagner, G., McLachlan, J., Pasquinelli, A.E., Balint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834-838.

Iliopoulos, D., Hirsch, H.A., and Struhl, K. (2009). An epigenetic switch involving NF-kappaB, Lin28, Let-7 MicroRNA, and IL6 links inflammation to cell transformation. Cell *139*, 693-706.

Jones, M.R., Quinton, L.J., Blahna, M.T., Neilson, J.R., Fu, S., Ivanov, A.R., Wolf, D.A., and Mizgerd, J.P. (2009). Zcchc11-dependent uridylation of microRNA directs cytokine expression. Nat Cell Biol *11*, 1157-1163.

Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S., and Baba, T. (2009). Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. Genes Dev *23*, 433-438.

Kwak, J.E., and Wickens, M. (2007). A family of poly(U) polymerases. RNA *13*, 860-867.

Lehrbach, N.J., Armisen, J., Lightfoot, H.L., Murfitt, K.J., Bugaut, A., Balasubramanian, S., and Miska, E.A. (2009). LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans. Nat Struct Mol Biol *16*, 1016-1020.

Loughlin, F.E., Gebert, L.F., Towbin, H., Brunschweiger, A., Hall, J., and Allain, F.H. (2011). Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28. Nat Struct Mol Biol.

Martin, G., and Keller, W. (2007). RNA-specific ribonucleotidyl transferases. RNA *13*, 1834-1849.

Martinez, N.J., and Gregory, R.I. (2010). MicroRNA gene regulatory pathways in the establishment and maintenance of ESC identity. Cell Stem Cell *7*, 31-35.

Melton, C., Judson, R.L., and Blelloch, R. (2010). Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. Nature *463*, 621-626.

Moss, E.G., Lee, R.C., and Ambros, V. (1997). The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA. Cell *88*, 637-646.

Nam, Y., Chen, C., Gregory, R.I., Chou, J.J., and Sliz, P. (2011). Molecular Basis for Interaction of let-7 MicroRNAs with Lin28. Cell *147*, 1080-1091.

Newman, M.A., Mani, V., and Hammond, S.M. (2011). Deep sequencing of microRNA precursors reveals extensive 3' end modification. RNA *17*, 1795-1803.

Newman, M.A., Thomson, J.M., and Hammond, S.M. (2008). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. RNA *14*, 1539-1549.

Piskounova, E., Polytarchou, C., Thornton, J.E., Lapierre, R.J., Pothoulakis, C., Hagan, J.P., Iliopoulos, D., and Gregory, R.I. (2011). Lin28A and Lin28B Inhibit let-7 MicroRNA Biogenesis by Distinct Mechanisms. Cell *147*, 1066-1079.

Piskounova, E., Viswanathan, S.R., Janas, M., LaPierre, R.J., Daley, G.Q., Sliz, P., and Gregory, R.I. (2008). Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28. J Biol Chem *283*, 21310-21314.

Rissland, O.S., Mikulasova, A., and Norbury, C.J. (2007). Efficient RNA polyuridylation by noncanonical poly(A) polymerases. Mol Cell Biol *27*, 3612-3624.

Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E.E., Nitsch, R., and Wulczyn, F.G. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. Nat Cell Biol *10*, 987-993.

Saitoh, S., Chabes, A., McDonald, W.H., Thelander, L., Yates, J.R., and Russell, P. (2002). Cid13 is a cytoplasmic poly(A) polymerase that regulates ribonucleotide reductase mRNA. Cell *109*, 563-573.

Schmidt, M.J., West, S., and Norbury, C.J. (2011). The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. RNA *17*, 39-44.

Siomi, H., and Siomi, M.C. (2010). Posttranscriptional regulation of microRNA biogenesis in animals. Mol Cell *38*, 323-332

Small, E.M., and Olson, E.N. (2011). Pervasive roles of microRNAs in cardiovascular biology. Nature *469*, 336-342.

Stagno, J., Aphasizheva, I., Bruystens, J., Luecke, H., and Aphasizhev, R. (2010). Structure of the mitochondrial editosome-like complex associated TUTase 1 reveals divergent mechanisms of UTP selection and domain organization. J Mol Biol *399*, 464-475.

van Wolfswinkel, J.C., Claycomb, J.M., Batista, P.J., Mello, C.C., Berezikov, E., and Ketting, R.F. (2009). CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. Cell *139*, 135-148.

Van Wynsberghe, P.M., Kai, Z.S., Massirer, K.B., Burton, V.H., Yeo, G.W., and Pasquinelli, A.E. (2011). LIN-28 co-transcriptionally binds primary let-7 to regulate miRNA maturation in Caenorhabditis elegans. Nat Struct Mol Biol *18*, 302-308.

Viswanathan, S.R., and Daley, G.Q. (2010). Lin28: A microRNA regulator with a macro role. Cell *140*, 445-449.

Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. Science *320*, 97-100.

Viswanathan, S.R., Powers, J.T., Einhorn, W., Hoshida, Y., Ng, T.L., Toffanin, S., O'Sullivan, M., Lu, J., Phillips, L.A., Lockhart, V.L.*, et al.* (2009). Lin28 promotes transformation and is associated with advanced human malignancies. Nat Genet *41*, 843-848.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855-862.

Winter, J., Jung, S., Keller, S., Gregory, R.I., and Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. Nat Cell Biol *11*, 228-234.
Wu, L., and Belasco, J.G. (2005). Micro-RNA regulation of the mammalian lin-28 gene during neuronal differentiation of embryonal carcinoma cells. Mol Cell Biol *25*, 9198-9208.

Wyman, S.K., Knouf, E.C., Parkin, R.K., Fritz, B.R., Lin, D.W., Dennis, L.M., Krouse, M.A., Webster, P.J., and Tewari, M. (2011). Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. Genome Res *21*, 1450-1461.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R.*, et al.* (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917-1920.

Zhu, H., Shyh-Chang, N., Segre, A.V., Shinoda, G., Shah, S.P., Einhorn, W.S., Takeuchi, A., Engreitz, J.M., Hagan, J.P., Kharas, M.G.*, et al.* (2011). The Lin28/let-7 axis regulates glucose metabolism. Cell *147*, 81-94.

CHAPTER 3

Sequence-specific mono-uridylation of mature miRNAs by Zcchc6 and Zcchc11

Text and figures partially adapted for an unpublished manuscript titled "Sequence-specific mono-uridylation of mature miRNAs by Zcchc6 and Zcchc11" by J.E. Thornton, P. Du, L. Sjekloca, E. Grossi, P. Sliz, and R.I. Gregory. The author is responsible for data in all figures except where noted. L. Sjekloca performed experiments in Figure 3.1D, E. Grossi performed Western blots in Figure 3.5, P. Du assisted in sequencing and bioinformatic analysis throughout. R.I. Gregory and P. Sliz assisted in designing experiments and were essential in discussions. The author and R.I. Gregory wrote the manuscript.

INTRODUCTION

MicroRNAs (miRNAs) are small non-coding RNAs that negatively regulate gene expression by binding to complementary sites in the 3' untranslated regions (3' UTRs) of messenger RNAs (mRNAs) and inducing gene silencing via the RNA-induced silencing complex (RISC) (Bartel, 2009). While the canonical pathway of miRNA biogenesis is largely understood, there are numerous types of modifications made to nascent miRNA species that impact their expression and function (Siomi & Siomi, 2010; Yates et al., 2013). Many of these modifications are appreciated in their ability to enhance or disrupt miRNA processing and repressive ability, while others are only recently being described and characterized (Burroughs et al., 2010; Chiang et al., 2010; Yates et al., 2013). Recently, several groups have uncovered the importance of widespread modification of 3' end of mature miRNAs and shown that this activity occurs for diverse miRNA species depending on the cell line, tissue, and developmental stage examined (Burroughs et al., 2010; Chiang et al., 2010; D'Ambrogio et al., 2012; Jones et al., 2012; Jones et al., 2009; Katoh et al., 2009; Wyman et al., 2011). Although the enzymes responsible for miRNA 3' nucleotide addition have in a few cases been identified—and their activities elucidated—much remains unknown about the nucleotide transferases involved in miRNA modification. In nearly every instance of mature miRNA modification only a subset of miRNAs are modified, but thus far no mechanism has been described to explain the underlying specificity of these reactions. Given that much of this activity is restricted to certain cellular contexts, the selective mechanisms of mature miRNA modification may play important roles in development and tissue specification.

RNA-dependent nucleotidyl transferases belong to the DNA Polymerase ß superfamily of enzymes and include the mRNA Poly(A) Polymerases (PAPs), CCA-adding enzymes

responsible for adding the 3' termini to transfer RNAs (tRNAs), and terminal uridyl transferases (TUTases) or poly(U) polymerases (PAPs) among others (Kwak & Wickens, 2007; Martin & Keller, 2007). There are seven mammalian RNA-dependent nucleotidyl transferases including Zcchc11 and Zcchc6 (TUTase4/TUT4/PAPD3 and TUTase7/TUT7/PAPD2, respectively), which play roles at several levels of miRNA biogenesis. Zcchc11 and Zcchc6 were initially identified as regulators of let-7 miRNA biogenesis in embryonic stem cells where they recognize the let-7 precursor RNA (pre-let-7) when it is bound by the small RNA-binding protein Lin28 (Thornton et al., 2012). The Lin28/pre-let-7/TUTase ternary complex is sufficient to oligo-uridylate the 3' end of pre-let-7 and facilitates its degradation by the exonuclease Dis3l2 (Chang et al., 2013; Heo et al., 2008). This miRNA turnover pathway serves to keep let-7 miRNAs low in undifferentiated cells and, upon the transition to differentiation, is relieved to ultimately permit the accumulation of mature let-7 miRNAs in differentiated cells (Hagan et al., 2009; Heo et al., 2009; Newman et al., 2008; Rybak et al., 2008; Thornton et al., 2012; Viswanathan et al., 2008). Zcchc11 and Zcchc6 along with GLD2/PAPD4, a TUTase previously shown to have mono- and oligo(A)-adding activity, together can add single uridine residues to the 3' end of certain precursor let-7 miRNAs to create a 3' end structure that facilitates efficient Dicer processing (Heo et al., 2012). This activity is Lin28-independent and is thought to occur constitutively in differentiated cells. Interestingly Zcchc11 was independently shown to uridylate the 3' end of mature miR-26 and thereby attenuate its activity towards target mRNAs involved in the inflammatory response (Jones et al., 2009). The repertoire of Zcchc11-targeted miRNAs was recently expanded to include a subset of mature miRNAs involved in growth hormone regulation and deletion of Zcchc11 disrupted normal growth patterns in fetal mice (Jones et al., 2012). It remains unclear how mono-uridylation blocks miRNA target repressive ability or how miR-26

and the several other miRNAs targeted by Zcchc11 and Zcchc6 are specified, but mature miRNA uridylation consistently fails to alter the steady-state levels of mature miRNAs and instead likely acts through changes in target specificity, RISC loading, or other mechanisms (Jones et al., 2012).

To further understand the role of Zcchc11 and Zcchc6 mature miRNA uridylation we investigated whether these two TUTases have an inherent preference towards various miRNA substrates and found that a simple sequence motif exists that is necessary and sufficient for targeting by Zcchc11 and Zcchc6 *in vitro*. Intriguingly, the only miRNAs that contain this motif are known regulators of developmental genes, and in several cases are highly enriched for targeting Hox gene 3'UTRs. The predicted miRNAs found in our search comprise four families, specifically: let-7, miR- 99/100, miR-196a/b, and miR-10a/b family members. In cultured cells the depletion of Zcchc11 and Zcchc6 led to a significant reduction in 3' uridylation in a small number of miRNA species, many of which could be predicted by our sequence-based approach. Furthermore we were surprised to find that after TUTase depletion and a corresponding drop in uridylated mature miRNAs there was a seemingly compensatory increase in mono-adenylation, effectively maintaining a constant pool of modified miRNAs. These results suggest that miRNA 3' uridylation is at least partially explained by sequence-specific activity of Zcchc11 and Zcchc6, and that this uridylation may impact the developmental program in animals. There may also be a miRNA PAP that functions in response to the loss of uridylation, and regulates the relative levels of modified versus unmodified miRNAs.

RESULTS

Sequence-specific miRNA uridylation by Zcchc11

Zcchc11 is known to be an important regulator of let-7 maturation in undifferentiated cells, yet its catalytic activity is not restricted to pre-miRNAs. In addition, several mature let-7 miRNAs are subjected to 3' mono- and oligo-uridylation but the mechanism underlying this activity is poorly understood (Burroughs et al., 2010; Chiang et al., 2010; D'Ambrogio et al., 2012; Hagan et al., 2009; Heo et al., 2012; Heo et al., 2009; Jones et al., 2009; Thornton et al., 2012; Wyman et al., 2011). To explore if Zcchc11 has preferential uridylation activity towards specific miRNA substrates we asked whether mature let-7 miRNAs can be targeted by immunopurified Zcchc11 *in vitro*. Flag-Zcchc11 purified from HEK293T cells was incubated in the presence of single stranded let-7g guide (let-7g-5p), let-7g passenger (let-7g* or let-7g-3p), or synthetic single stranded RNAs designed to target luciferase (GL2 guide and passenger). Using equal amounts of RNA as determined by 5' end-labeling, we performed *in vitro* uridylation assays and monitored TUTase activity by the incorporation of radiolabeled UTP. We found Zcchc11 preferentially uridylates let-7g guide over let-7g passenger or either strand of luciferase-targeting small RNAs (Figure 3.1A). We also found Zcchc11 to preferentially recognize single-stranded RNA (ssRNA) over double-stranded RNA (dsRNA) as described previously (Jones et al., 2009) (data not shown).

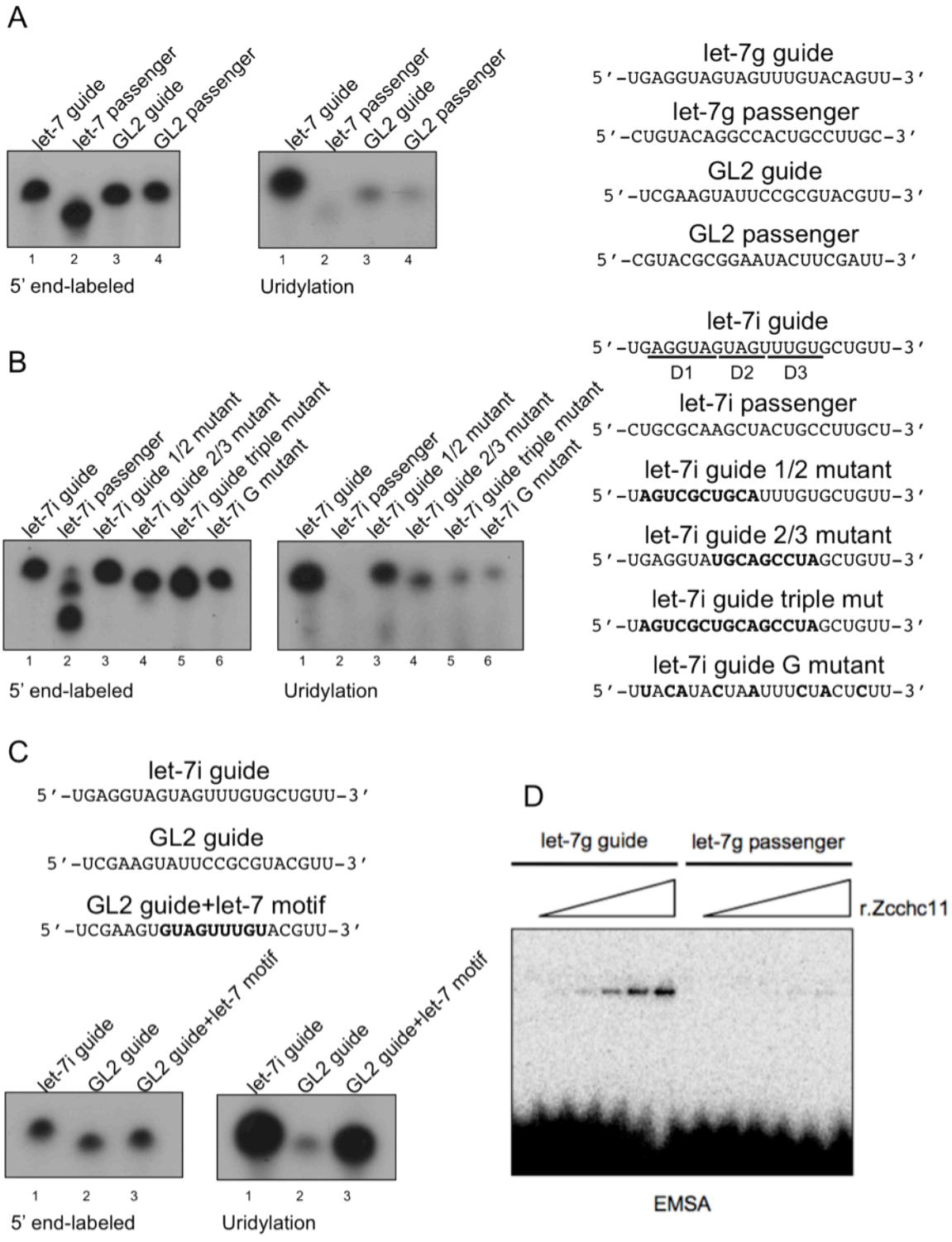


Figure 3.1: Zcchc11 uridylates mature miRNAs in a sequence dependent manner

Figure 3.1 continued

A. *In vitro* uridylation assays performed with different synthetic RNA substrates reveal preferential Zcchc11 uridylation activity for let-7 miRNAs over non-let-7 miRNAs.

B. A sequence motif in the central region of let-7i guide miRNA is necessary for Zcchc11 preferential activity *in vitro*.

C. A let-7 sequence motif found in let-7i and let-7g is sufficient to confer preferential Zcchc11-mediated uridylation activity *in vitro*.

D. EMSA with the indicated radiolabeled RNA probes (2.5 nM final concentration) and increasing amounts (1x, 2x, 4x, 8x, and 16x) of recombinant Zcchc11 protein.

The dramatic uridylation preference for one strand of let-7g duplex over the other suggested the presence of specific sequences in let-7g guide that were absent in let-7g passenger. To determine the sequences that convey this substrate specificity we performed mutagenesis on let-7i guide, another member of the let-7 family. Three contiguous regions of the let-7i guide strand were chosen for further examination; the seed sequence (domain 1), a GXXG motif in the center of the RNA (domain 2) since it is known that CCHC zinc-finger proteins can bind to GXXG motifs (Heo et al., 2009; Loughlin et al., 2011; Nam et al., 2011), and a U-rich region near the 3' end of let-7i guide (motif 3). Consistent with the preference shown towards let-7g, let-7i guide was subjected to stronger uridylation activity than let-7i passenger, whereas individually mutagenized RNA for each of the three domains showed no remarkable reduction in uridylation (Figure 3.1B and data not shown). Interestingly, when combinatorial mutations were made in domains 2 and 3 (Δ2/3), uridylation was more substantially compromised than mutations in domains 1 and 2 (Δ1/3). Mutations in all three domains (Δ1/2/3) showed a reduction in uridylation similar to the domain 2/3 double mutant, implying that this region is necessary for sequence-specific uridylation. Finally, a let-7i RNA oligo lacking all guanosine residues (ΔG) was tested, since guanosine-rich sequences are known to mediate the binding of zinc finger-RNA interactions (Heo et al., 2009; Loughlin et al., 2011; Nam et al., 2011). This mutant also led to compromised activity, similar in extent to the domain 2/3 double mutant and the triple mutant, suggesting that guanosine residues play a critical role in Zcchc11 substrate recognition.

To examine whether domains 2 and 3 of let-7i are sufficient for driving enhanced uridylation, a synthetic RNA designed to target luciferase was modified to contain these regions. Uridylation activity towards GL2 was much weaker compared to let-7i, but was almost completely restored when the let-7 motif was inserted into the GL2 RNA oligo (Figure 3.1C).

These data suggest that specific sequences found in several let-7 miRNAs are both necessary and sufficient to drive preferential uridylation activity of Zcchc11.

We also measured the relative binding of recombinant Zcchc11 to either the let-7 guide or let-7 passenger RNA sequences. Consistent with the preferential uridylation activity we observed, we found purified Zcchc11 preferentially associates with let-7 guide RNA in an electrophoretic mobility shift assay (EMSA) (figure 3.1D).

The sequence motif conferring Zcchc11-catalyzed uridylation is present in a small subset of miRNAs

Several mature miRNAs have recently been described that contain non-templated 3' U residues (Burroughs et al., 2010; Jones et al., 2012; Jones et al., 2009; Wyman et al., 2011), with a limited number undergoing a reduction in uridylation after Zcchc11 depletion (Jones et al., 2012; Wyman et al., 2011). To gain insight into the mechanism of substrate selectivity and to identify other miRNAs that contain a similar uridylation signal, we performed an alignment of all members of the human let-7 miRNA family. As shown in Figure 3.2A, all let-7 members contain at least one GUAG sequence, while all except let-7e and miR-98 contain two overlapping GUAG motifs. Surprisingly, however, only let-7i and let-7g contain the UUUGU motif identified above. The first nucleotide in this sequence (shown in red) is poorly conserved relative to the remaining sequences shared by let-7 miRNAs. We predicted that motifs involved in mature miRNA regulation should be found in all or most members of a given miRNA family, and considered

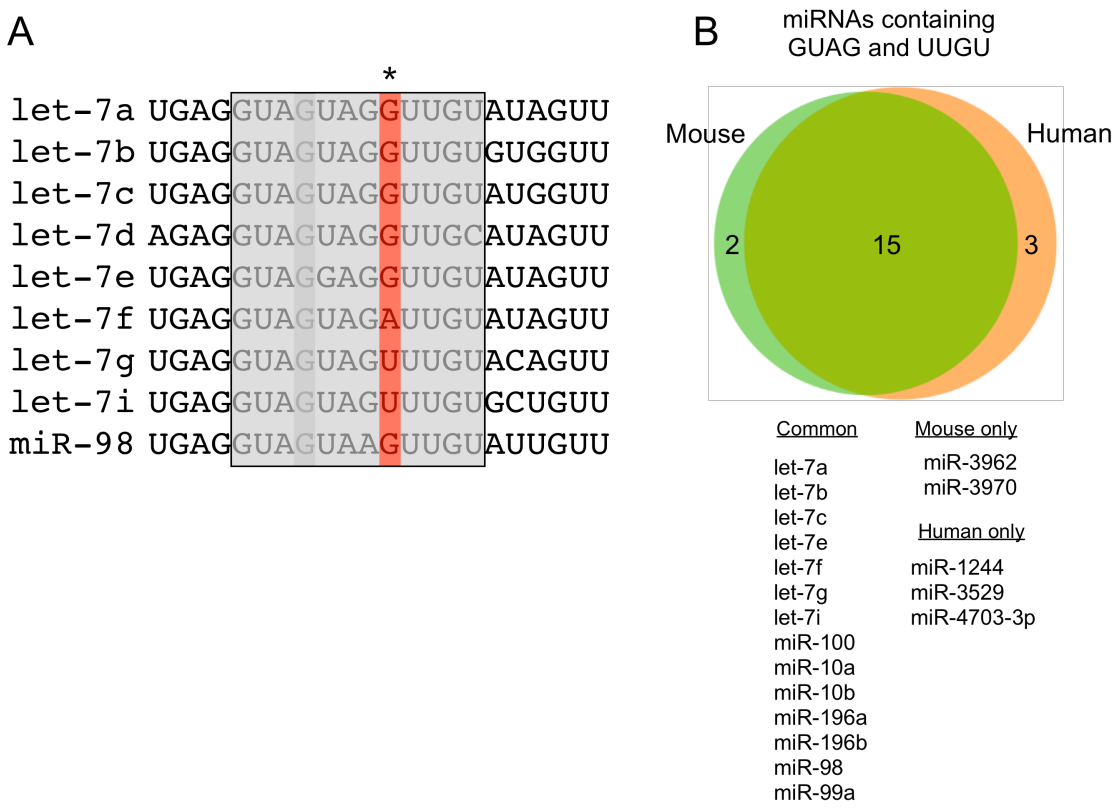


Figure 3.2: A motif in mature let-7 miRNAs defines Zcchc11-targeted substrates

Figure 3.2 continued

A. A uridine residue identified as part of the let-7 motif sequence is poorly conserved and may be dispensable for mature let-7 uridylation.

B. Searching mouse and human miRNA databases reveals a highly overlapping subset of mature miRNAs containing the Zcchc11-targeting motif.

C. miR-10a contains the let-7 targeting motif and is a substrate for Zcchc11 *in vitro* while miR-26, which has previously been described as a Zcchc11 substrate in cultured cells lacks the intrinsic Zcchc11-targeting property and is a poor substrate as seen from *in vitro* uridylation assays.

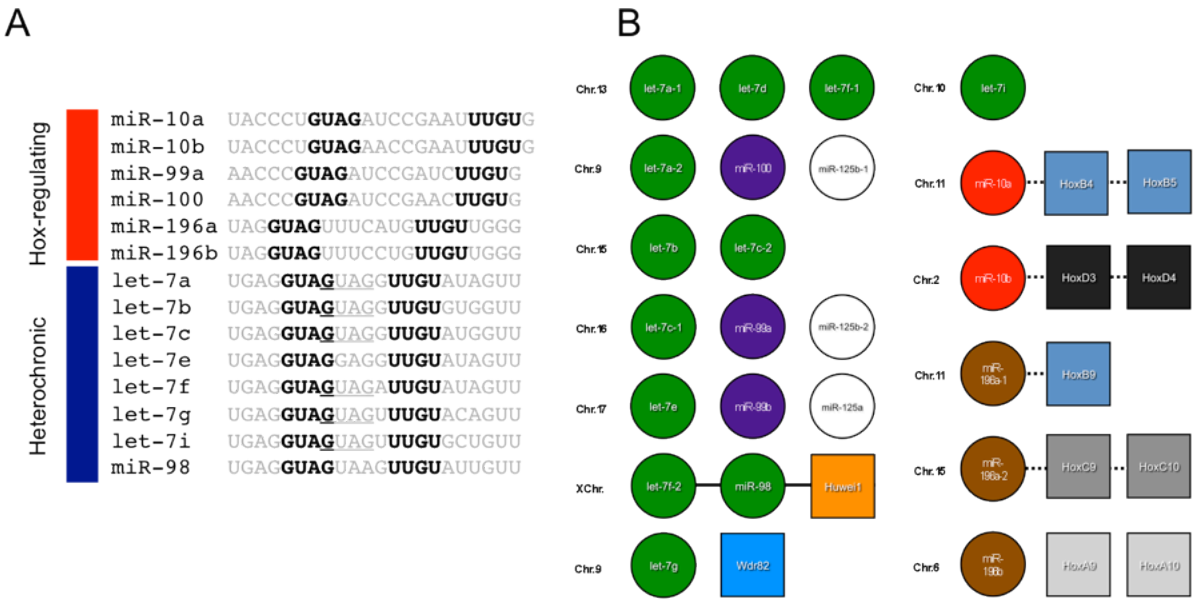
that the domain 2/3 motif may be too stringent and result in false negatives. In light of this, we performed further bioinformatics analyses to determine miRNAs that contain the more limited motif of GUAG and UUGU. The reduced stringency of this search could potentially introduce sampling bias and lead us to overestimate putative Zcchc11 target miRNAs, so we compared the lists of miRNAs containing these two sequences in mouse and human databases to observe whether we were identifying a large number of species-specific miRNAs. Importantly, nearly every miRNA identified in this search is found in both mouse (15/17) and human (15/18) databases, while those miRNAs restricted to one species or the other have been identified exclusively in high-throughput sequencing experiments, have no known biological function, and often lack homologs in closely related species (Figure 3.2B).

To determine if the set of putative Zcchc11 substrate miRNAs predicted above are preferred *in vitro*, we carried out uridylation assays using Flag-Zcchc11 and a synthetic RNA oligo for one of these candidates, miR-10a guide (miR-10a-5p). Comparing the activity of Zcchc11 towards let-7i guide, a similar level of uridylation was observed for miR-10a but was lost when the GUAG and UUUGU motifs were mutated to random nucleotides (Figure 3.2C, lower panel). We also tested the intrinsic preference of Zcchc11 towards miR-26a, a known substrate, but found that it did not support strong uridylation similar to let-7 or miR-10. Overall these data support the idea that the identified domains (GUAG and UUGU) are sufficient to predict selective monouridylation by Zcchc11 *in vitro*.

TUTase motif-containing miRNAs define a family of developmental regulators

Examining the list of predicted Zcchc11 miRNA substrates led to a striking observation; all of the miRNAs in this analysis fall into the broad category of development regulatory

miRNAs (Figure 3.3A). Roughly half of this group, the let-7 miRNA family, is the best-characterized heterochronic miRNA family, known to regulate developmental timing in organisms as diverse as *C. elegans* and humans (Pasquinelli et al., 2000; Thornton & Gregory, 2012). The remaining miRNAs found in our search prediction are known regulators of development, functioning primarily through repressing Hox genes during posterior-anterior patterning in vertebrates (Yekta et al., 2008). The three families we found, miR-10, miR-99/100, and miR-196, are highly conserved and are dysregulated in a number of human diseases. Intriguingly, many of these miRNAs are expressed from the same genomic clusters and are often encoded within various Hox loci (Figure 3.3B). To examine the relative importance of these three miRNA families in Hox gene regulation, we asked which miRNA families most frequently target the 3'UTR of Hox gene mRNAs. In both human and mouse genomes, these three miRNA families are some of the most likely to target Hox gene mRNAs, and several Hox gene UTRs contain multiple sites for several different members of these miRNA families (Figure 3.3C and data not shown). For the various members of the predicted miRNA families, this sequence-based approach identified all members except for let-7d and miR-99b. In both instances these miRNAs were not predicted based on a U to C transposition at the final position of UUGU. The enrichment of Hox-targeting and heterochronic miRNAs suggests that Zcchc11 may specifically target developmental miRNAs.



C

	Human			Mouse		
	Total Conserved Target Sites	Hox Targets (Percent)	Rank (out of 78 families)	Total Conserved Target Sites	Hox Targets (Percent)	Rank (out of 69 families)
miR-196	**328**	**14** (4.27%)	**1**	**238**	**7** (2.94%)	**1**
miR-99/100	**59**	**2** (3.39%)	**2**	**37**	**1** (2.70%)	**2**
miR-10	**300**	**6** (2.00%)	**3**	**207**	**5** (2.42%)	**4***

***miR-335 ranks 3rd at 2.63%**

Figure 3.3: Zcchc11 motif-containing miRNAs are all involved in development

Figure 3.3 continued

A. Schematic representation indicating the regions that define a Zcchc11-targeting motif (shown in bold and underlined). Two classes of miRNAs are defined by this sequence, namely the Hox-gene targeted miRNAs and the let-7 family of miRNAs.

B. Many of the miRNAs predicted to have Zcchc11-targeting sequences are co-expressed in transcriptional clusters or from within Hox gene loci.

C. Non-let-7 family members described in Fig 3A are highly predicted to target Hox genes in mice and humans.

TUTase domains required for selective miRNA uridylation

Zcchc6 is a TUTase that shares significant homology and redundant activity with Zcchc11 in regulating pre-let-7 levels in embryonic stem cells (Thornton et al., 2012) (Figure 3.4A). We tested whether this redundant activity was also shared for the ability to uridylate a subclass of mature miRNAs (Figure 3.4B). Consistent with previous work showing similar activity between the two proteins, we found Zcchc6 to have as strong a preference for let-7 guide and miR-10a guide over let-7 passenger and miR-26 as we did for Zcchc11 (Figure 3.4B). The full-length Zcchc11 is an 184kDa non- canonical poly(A) polymerase that is highly conserved across vertebrates. Zcchc6 is slightly smaller and lacks domains at the N- and C- termini. The Zcchc6/11 active site is located within the Nucleotidyl Transferase (Ntr) domain, which is paired with a Poly(A)-Polymerase-Associated (PAP) domain; a common feature of non-canonical poly(A) polymerases (Kwak & Wickens, 2007; Lunde et al., 2012; Martin & Keller, 2007; Munoz-Tello et al., 2012; Saitoh et al., 2002; Yates et al., 2012). Catalysis requires a conserved aspartate triad in the Ntr (Hagan et al., 2009). Flanking the active site are three CCHC retroviral-type zinc fingers/zinc knuckles. At the N-terminus of the protein is a region that shares significant homology with the active site, including a proximal PAP domain; however, this region lacks one of the crucial aspartates predicted to be necessary for catalysis. Instead, this region is most similar to the yeast TRF4 proteins, which carry out cytoplasmic poly(A) RNA polymerase activity (Saitoh et al., 2002). N-terminal to this region is a classical C2H2 zinc finger that we previously found to be required for functional interaction with Lin28 for the enhanced oligouridylation of pre-let-7 RNA (Thornton et al., 2012) (Figure 3.4A). To understand which domains of Zcchc11 are required for mature miRNA uridylation activity, we generated mutant

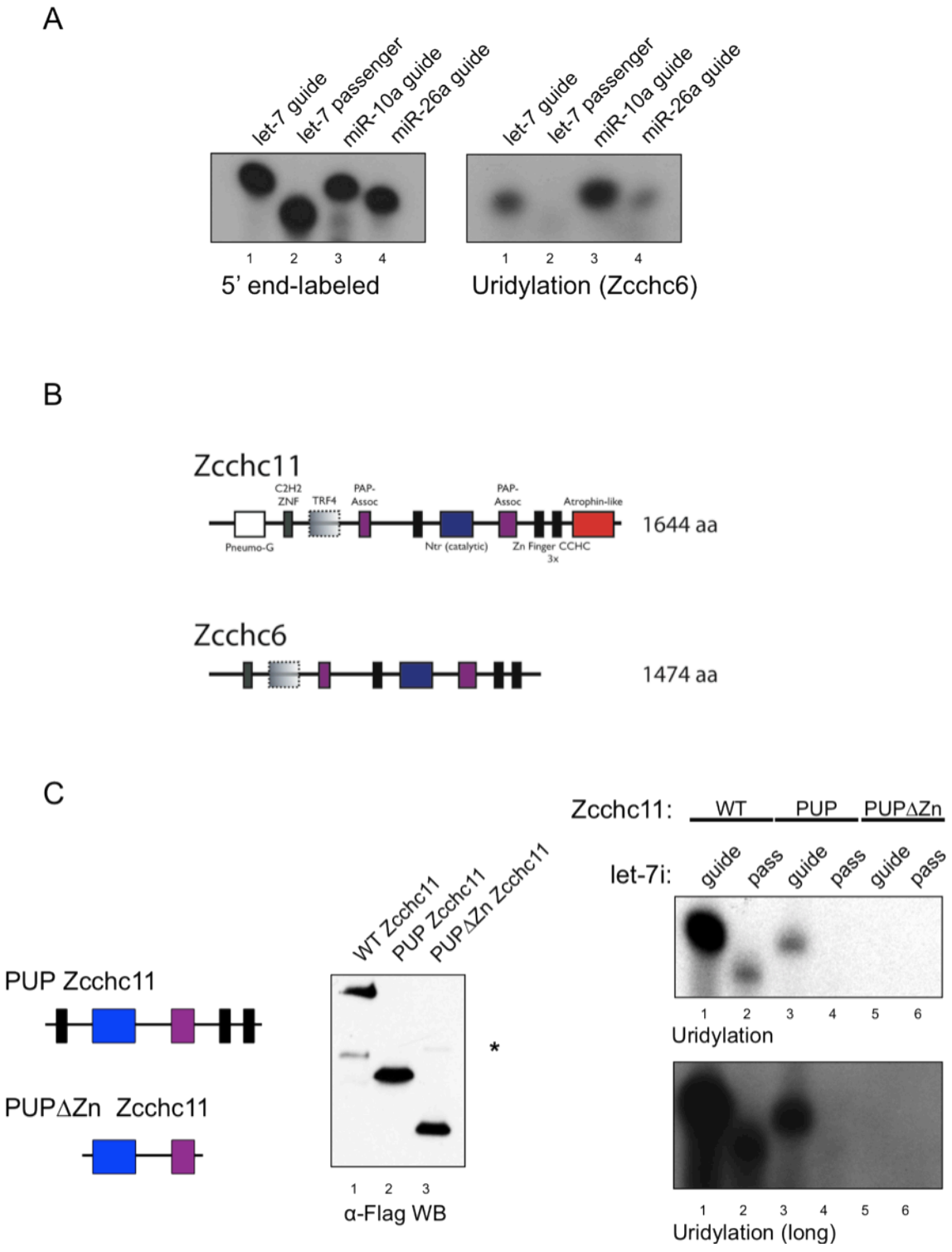


Figure 3.4: TUTase domains required for selective miRNA uridylation

Figure 3.4 continued

A. Schematic representation of mouse Zcchc6 and Zcchc11 proteins.

B. Zcchc6 selectively targets miRNAs *in vitro*.

C. The CCHC zinc fingers of Zcchc11 are required for robust uridylation activity but the N- and C-termini of this protein are partially dispensable for selective mature miRNA uridylation. Flag western blot of the indicated proteins (left) and *in vitro* uridylation assays with the indicated synthetic RNA substrates (right).

cDNAs and tested the ability of the resulting Flag-immunopurified (Flag IP) proteins to uridylate let-7 miRNA *in vitro* (Figure 3.4C). Mutants were generated lacking N- and C-terminal domains alone (PUP) or lacking N- and C-terminal domains plus the CCHC zinc fingers (PUPΔZn) (Figure 3.4C). Although the overall activity of the PUP mutant was found to be less than the full-length protein, this truncated protein retained the ability to selectively uridylate let-7 guide RNA. Loss of the Zinc fingers rendered the protein inactive, indicating that the CCHC zinc finger domain(s) may be essential for the interaction with RNA. These studies provide insight into the basic mechanism underlying the catalytic nature of Zcchc11 towards miRNA substrates (Figure 3.4C).

Zcchc6 and Zcchc11 expression is developmentally regulated

The extent of 3' terminal uridylation of individual miRNAs is reported to be variable during cell differentiation, development, and in different tissues (Chiang et al., 2010; Wyman et al., 2011). Considering these features we next examined the expression of Zcchc6 and Zcchc11 proteins in a panel of different adult mouse tissues where let-7 miRNAs are abundantly expressed. Interestingly, this analysis revealed that expression of both of these TUTases is undetectable in most adult tissues examined (Figure 3.5A). These results are consistent with the recently reported age-dependent expression of Zcchc11, with strongest expression observed in most organs at young ages (Jones et al., 2012). Extract from P19 embryonal carcinoma (EC) cells was used as positive control since we have previously reported an overlapping function for Zcchc6/11 in the Lin28-mediated control of let-7 biogenesis in these cells (Thornton et al., 2012). This raised the possibility that expression of Zcchc6 and Zcchc11 is dynamically regulated during development and/or cell differentiation. To further investigate this we treated

P19 cells with retinoic acid and monitored Zcchc6/11 expression as cells differentiated over a time-course of several days. We found that while Zcchc6 and Zcchc11 protein was robustly detected in undifferentiated cells the expression of these proteins was rapidly downregulated during cell differentiation (Figure 3.5B).

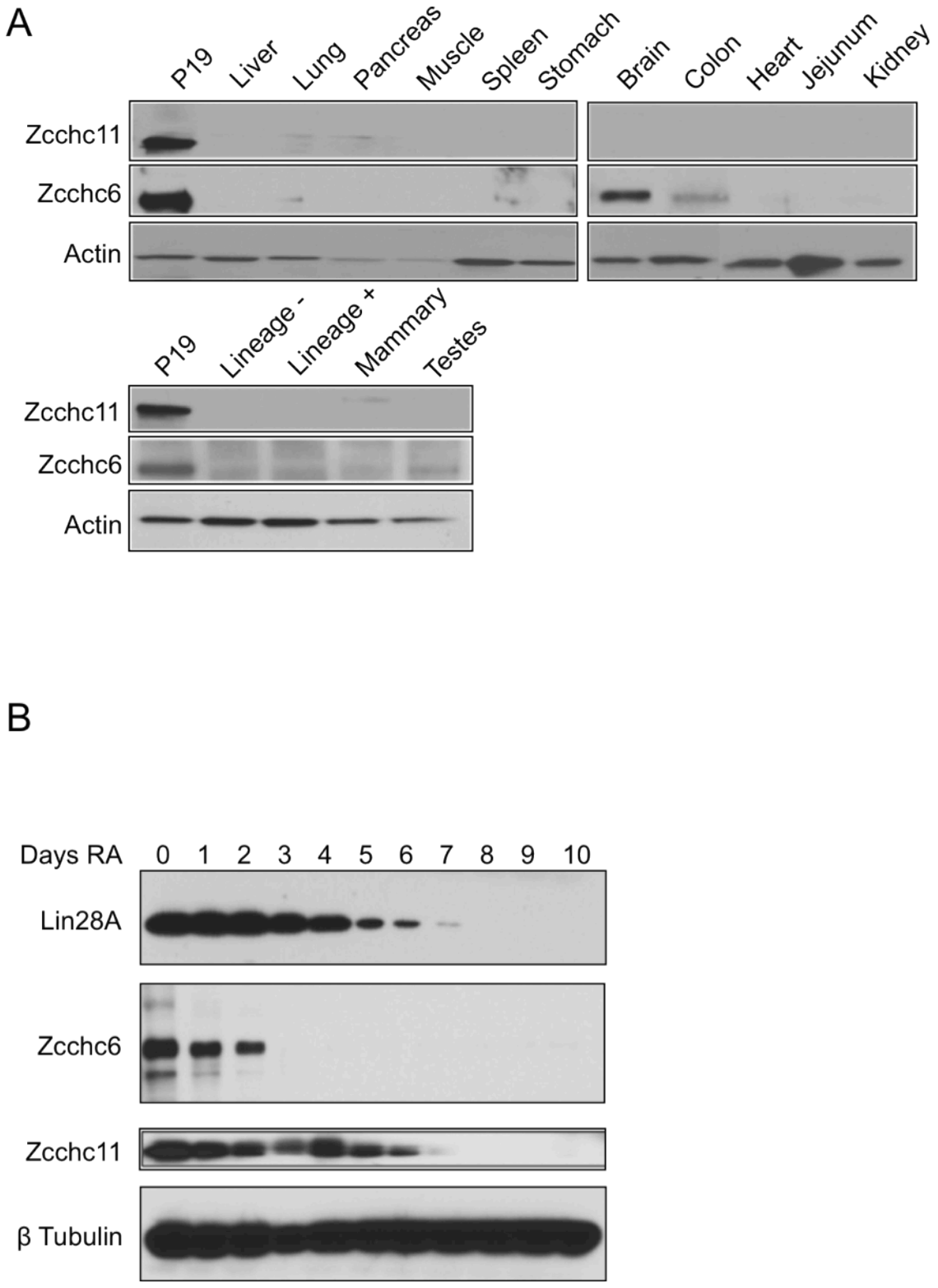


Figure 3.5: Zcchc11 and Zcchc6 are developmentally regulated TUTases

Figure 3.5 continued

A) Zcchc11 and Zcchc6 are largely absent from differentiated tissues with the exception of Zcchc6 in the brain and colon.

B) Zcchc11 and Zcchc6 are highly expressed in embryonic stem cells but are rapidly down-regulated during retinoic acid-mediated differentiation with kinetics either the same as or more rapidly than the pluripotency factor Lin28.

Depletion of Zcchc6 and Zcchc11 does not affect miRNA abundance

Our biochemical assays suggest that Zcchc11 and Zcchc6 may regulate the uridylation of a specific subset of miRNAs. To test this hypothesis we examined the impact of Zcchc11 and Zcchc6 depletion on miRNA levels in cultured cells. Because both Zcchc11 and Zcchc6 function with Lin28 to regulate pre-let-7 in many cell types, we performed the following experiments in Hela cells, as they do not express either Lin28A or Lin28B, but express both TUTases (Figure 3.6A) (Piskounova et al., 2011). Hela cells were stably transduced with lentiviruses targeting Zcchc11, Zcchc6, or both genes together using different selectable markers (Figure 3.6B). Total RNA was purified from these stable cell lines and levels of let-7g were analyzed by Northern blot. There was no detectable change in the abundance of let-7g, consistent with previous reports that showed no change in the steady-state levels of mature miRNAs after Zcchc11 depletion (Jones et al., 2009; Jones et al., 2012). Interestingly, the minor let-7g species detectable by Northern blot did not change in either, suggesting that these bands may arise from cleavage heterogeneity rather than non-templated nucleotide addition (Figure 3.6C). To more comprehensively examine the overall levels and 3' status of mature miRNAs in TUTase-depleted Hela cells, we cloned and sequenced small RNA libraries from control (shGFP) and Zcchc11/Zcchc6 double knockdown cells (shTUT). Total reads of all miRNAs irrespective of 3' status were examined and remained largely unchanged (Figure 3.6D). The relative wild-type levels of predicted TUTase substrate miRNAs were also largely unchanged, with no changes larger than 2-fold in either direction (Figure 3.6E). These experiments indicate that depletion of Zcchc11 and Zcchc6 do not dramatically change the expression landscape of miRNAs in Hela cells, even those predicted as substrates of these two TUTases.

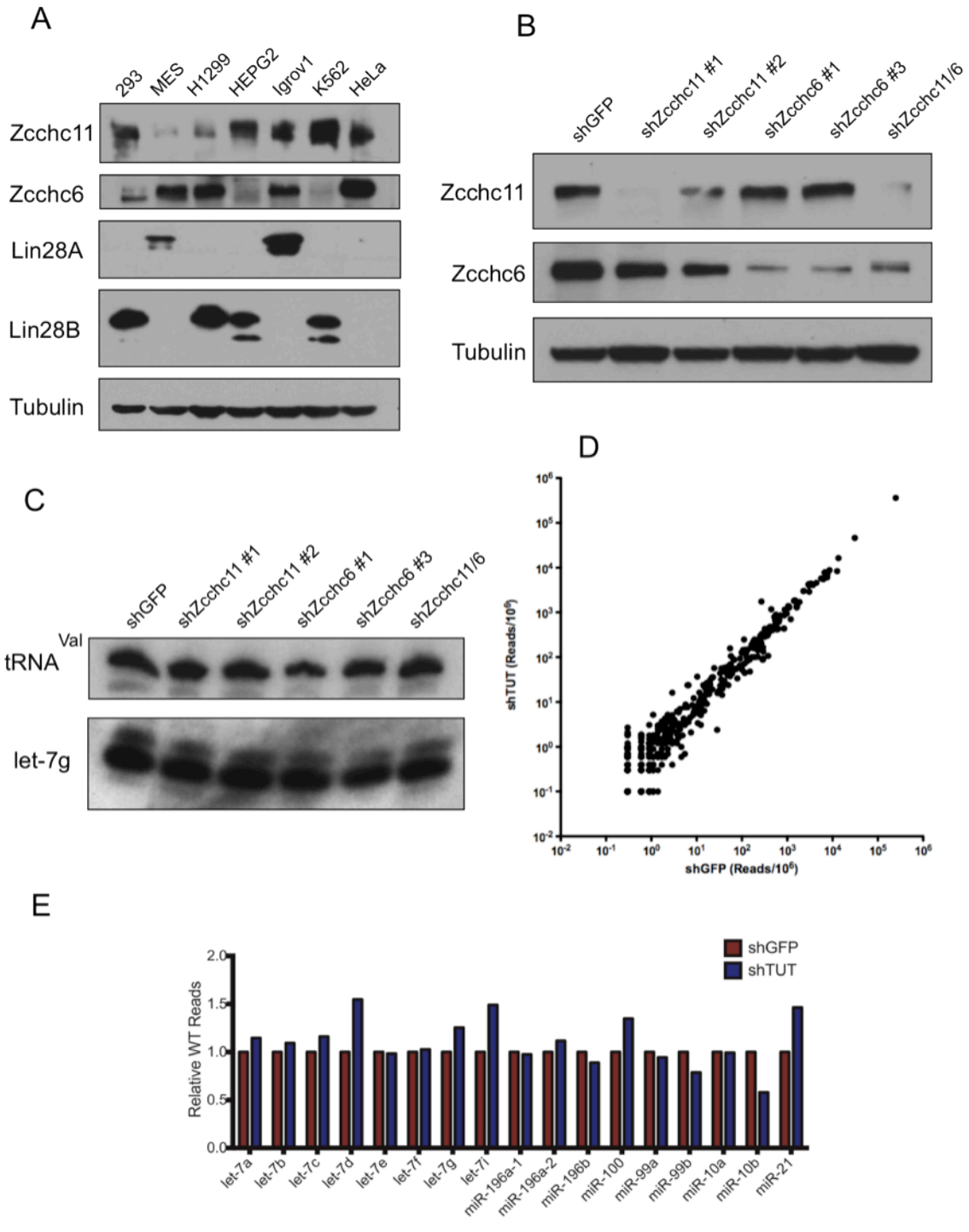


Figure 3.6. TUTase depletion does not affect miRNA levels

Figure 3.6 continued

A. Hela cells were express neither Lin28A nor Lin28B but have high expression of both Zcchc11 and Zcchc6.

B. Letnivirus-mediated knockdown of Zcchc11 and Zcchc6 in Hela cells.

C. Mature let-7g levels do not change in the absent of TUTases as determined by Northern blotting.

D. TUTase depletion does not affect global mature miRNA levels as determined by small RNA sequencing.

E. Sequences of motif-containing mature miRNAs are largely unchanged after TUTase depletion.

TUTase depletion causes selective loss of miRNA mono-uridylation and concomitant gain of mono-adenylation

To determine if the 3' status of miRNAs predicted in this study are dramatically altered in the absence of Zcchc11 and Zcchc6, we compared their overall uridylation status with a large subset of mature miRNAs expressed in Hela cells. One major obstacle in studying mature miRNA uridylation is determining at which processing step non-templated nucleotides are added, since mature miRNAs can be produced from either side of a pre-miRNA hairpin. If mature miRNAs are derived from the 3' arm of the precursor loop (3p-miRs), then uridylation can occur upstream of Dicer. These 3p-miRs may also undergo mature miRNA uridylation after Dicer cleavage or after strand selection and incorporation into RISC. Conversely, mature miRNAs derived from the 5' arm of a precursor loop (5p-miRs) can only undergo uridylation after Dicer cleavage, either before or after strand selection and RISC incorporation. Furthermore, non-templated U addition is enriched for 3p-miRs and widespread pre-miRNA uridylation has been reported for several distinct miRNA families beyond let-7, complicating the analysis of 3p-miR uridylation (Chiang et al., 2010; Newman et al., 2011). Fortuitously, all of the predicted miRNAs from our analysis are derived from the 5p arm of their precursor hairpins as determined by the relative read number in our experiments as well as by online databases such as miRBase (Griffiths-Jones et al., 2006). To examine the levels of uridylation for predicted TUTase substrate miRNAs, we compared the 3' ends of these miRNAs to all other 5p-miRs so that we were exclusively observing mature miRNA uridylation rather than a combination or pre- miRNA and mature miRNA effects. For a more stringent comparison of 5p-miRs we examined only

those that contained non-templated uridylated residues at a frequency of 10 reads per million (RPM) and made up no less than 1% of reads for a single miRNA in the shGFP sample. Of the 19 miRNAs that fall into this category in shGFP Hela cells, uridylation occurs between 1% and 8% and in all cases it is substantially reduced after Zcchc11 and Zcchc6 depletion (Figure 3.7A). Notably, several of the miRNAs predicted in our analysis appear in this group, making up 5 of the top 10 most-uridylated miRNAs (Figure 3.7A, underlined). Several predicted miRNAs fell just below the threshold criteria, including miR-196b, let-7e, miR-98, and let-7b. In all instances these miRNAs also underwent a significant loss of uridylation similar to those listed above (data not shown).

In addition to ranking the most-uridylated miRNAs, we wished to examine the composition of the 3' ends of this miRNA subclass. As shown in Figure 3.7B, miR-99a predominantly exists in its 22-nucleotide (nt) form, with a significant proportion being 21nt in length. The next most common species is the 22nt isoform extended by a nontemplated adenosine residue, followed by the 22nt form extended by a single uridine residue. Reads containing a non-templated uridine residue are significantly reduced after TUTase depletion, however we observe an increase in adenylated reads in response to TUTase depletion as well. It may also be the case that the cleavage site for miR-99a production is shifted from the 22nt to the 21nt form after TUTase loss. This could be due to changes in 3' pre-miRNA uridylation altering the Dicer cleavage site, or this could be due to other unexplored effects of mature miRNA uridylation.

To understand the 3' modification landscape of modified miRNAs, we compared the terminal nucleotides of these miRNAs in shGFP and shTUT cells. Restricting our analysis to sequences that are present at 1RPM or greater in shGFP cells, genomic reads of miR-99a make

up 80% of all reads, with non-templated uridine residues contributing 8%, non-templated adenosines contributing 11%, and the remainder coming from cytidine or guanine addition. Intriguingly, as uridylation levels are decreased after knockdown of Zcchc11 and Zcchc6, there is a concomitant increase in non-templated adenosine residues (Figure 3.7C). This phenomenon is found in nearly all predicted TUTase substrate miRNAs and is common in nearly all of the top-uridylated miRNAs (Figure 3.7D). The impact of this compensatory adenylation is unknown and it remains unclear which enzyme is responsible for this modification. These experiments together show that a select group of miRNAs undergo Zcchc11 and Zcchc6-dependent uridylation in cultured cells, and that half of the top miRNAs undergoing uridylation can be predicted by a simple sequence motif from *in vitro* studies. Furthermore, with the loss of the these two TUTases and the marks of their catalytic activity, non-templated adenylation often appears, which effectively retains a constant proportion of miRNA reads corresponding to genomic sequences.

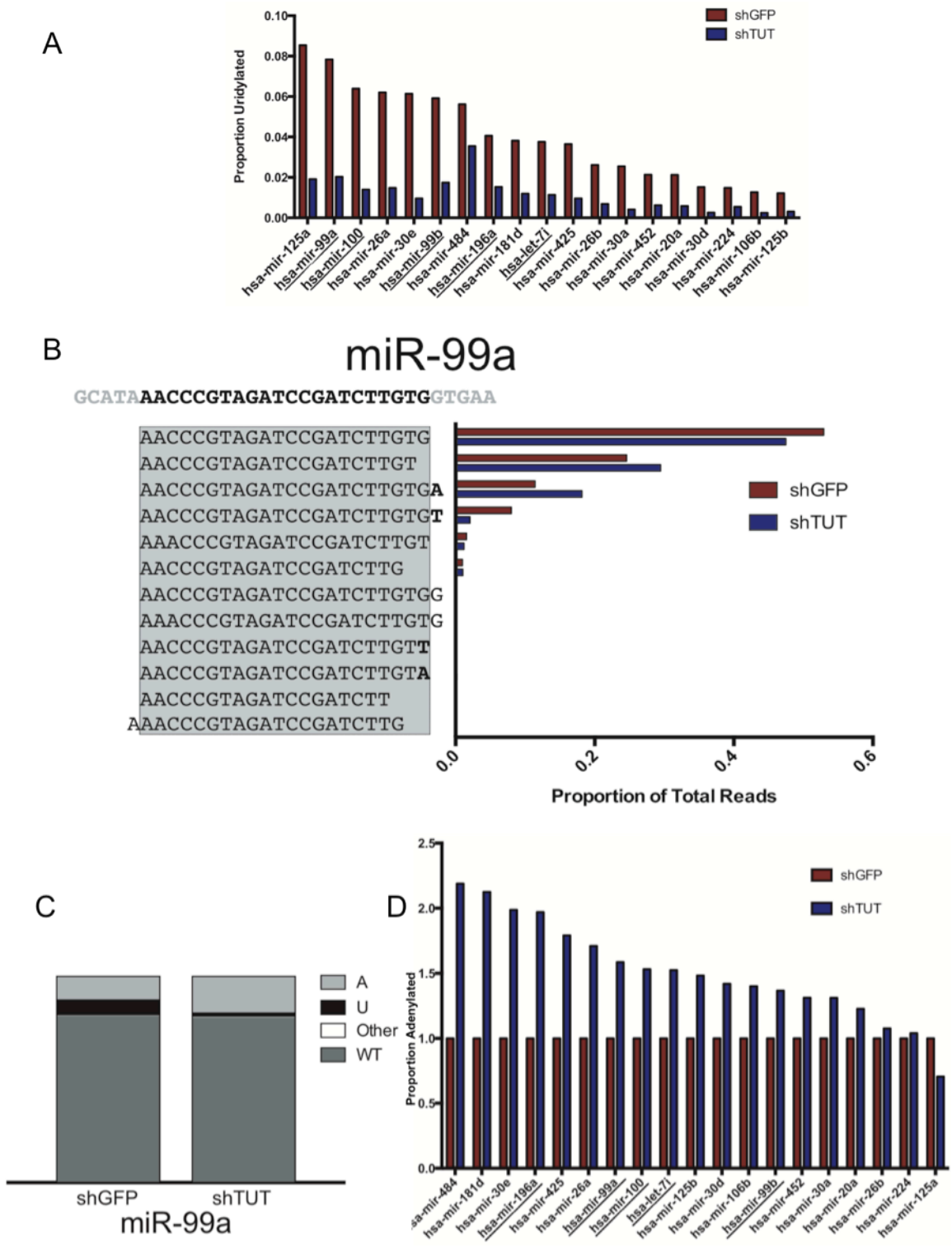


Figure 3.7: Mature miRNAs are uridylated in cultured cells in a TUTase-dependent manner

Figure 3.7 continued

A. The most highly uridylated miRNAs are ranked as according to 3' non-templated uridine status. Underlined miRNAs were predicted by our sequence search.

B. Mature miRNA uridylation leads to the reduction of reads containing non-templated uridine residues and leads to the accumulation of adenylated miRNAs. The genomic sequence of the miR-99a locus is listed at top, the gray box at left represents the canonical mature miRNA sequence, and the bold residues are nontemplated nucleotides added after Dicer cleavage.

C. Distribution of 3' terminal residues for miR-99a in control and TUTase-depleted cells.

D. After TUTase depletion the most highly uridylated miRNAs are in many cases subjected to 3' adenylation.

DISCUSSION

In the last several years a number of activities have been credited to Zcchc11 and Zcchc6 including positive and negative regulation of pre-let-7 miRNAs, mediating toll-like receptor (TLR) signaling, uridylation of a small subset of mature miRNAs, and regulating the cell cycle independent of its catalytic activity (Blahna et al., 2011; Hagan et al., 2009; Heo et al., 2012; Heo et al., 2009; Jones et al., 2012; Jones et al., 2009; Minoda et al., 2006; Piskounova et al., 2011; Rau et al., 2011; Schmidt et al., 2011; Thornton et al., 2012). Our findings add to the existing catalog of TUTase substrate RNAs and are the first examples to our knowledge of a sequence-dependent RNA-protein interaction for either of these enzymes. This specific activity is reminiscent of the U6 TUTase (TUTase6/TUT1/PAPD2/Hs5), which adds an oligo-uridine tail exclusively to the U6 snRNA (Trippe et al., 2006). Mature miRNA uridylation by Zcchc11 and Zcchc6 has been shown to disrupt target mRNA repression without causing a change in the steady state levels of the miRNA (Jones et al., 2012; Jones et al., 2009). It remains unclear how only one or two non-templated nucleotides can disrupt targeting of mRNAs since the 3' end of miRNAs is thought to be largely dispensable for repression by RISC (Wee et al., 2012). This question warrants further investigation and remains one of the most outstanding questions in the field.

Upon identifying a sequence motif that is necessary and sufficient for miRNA targeting by Zcchc11 and Zcchc6, we were surprised to find that only developmental miRNAs contained this sequence, notably, miRNAs that target Hox genes. The miR-10, miR-99/100, and miR-196 families are some of the most common Hox gene-targeting miRNAs yet described. They are expressed either in clusters containing let-7 family members or from within Hox gene clusters

themselves, and are known to regulate numerous Hox genes—in many cases those located in a nearby genomic locus. Hox-targeting by these miRNAs is thought to repress anterior expression patterns and promote posterior development (Yekta et al., 2008). One notable example is the ability for miR-196a to bind a perfectly complementary 22nt sequence in the 3'UTR of HoxB8, a rare case in mammals that causes cleavage of the transcript similar to the activity of an siRNA (Hornstein et al., 2005; Wee et al., 2012). Given the known ability for uridylation to attenuate miRNA repression of target genes, it is possible that temporal regulation of Hox-targeting miRNAs may play a role in embryonic patterning. Although a Zcchc11 knockout mouse has been described and survives to birth at Mendelian ratios, the redundancy between the two TUTases in mature miRNA uridylation and pre-let-7 turnover suggests a double knockout animal is required before robust phenotypes are observed (Jones et al., 2012).

Interestingly, miR-125 family members are expressed from some of these same genomic clusters but do not contain the predicted targeting sequence. Still, they are highly uridylated in cultured Hela cells and undergo a robust decrease in uridylation after loss of Zcchc11 and Zcchc6, and this is particularly true for miR-125a which was the most highly uridylated mature miRNA in our study. It is tempting to speculate that the miR-125 family—the mammalian homolog of the *C. elegans* heterochronic miRNA lin-4—is also targeted by Zcchc11 and Zcchc6 due to its known role in development, yet the mechanism of this targeting remains unknown. Indeed, we found miR-26 family members to undergo TUTase-dependent uridylation as reported previously even in the absence of the sequence motif, indicating that there are other levels of specificity that remain unknown. Also worth noting is the similarity in the seed sequence between let-7 family miRNAs and miR-196a/b. Although they belong to distinct miRNA families their seed sequences are different only in that they are shifted by one nucleotide,

suggesting that there may be an overlap in target mRNAs for these two miRNA families.

Sequencing of small RNAs isolated from shTUT Hela cells uncovered uridylated reads for many of the miRNAs predicted by our *in vitro* studies. One complication of this approach is the ambiguity that arises when the genomic sequence of a miRNA is such that nucleotides after the canonical 3' end of 5p-miRNAs are thymidine-rich. In these instances it is impossible to determine if a terminal uridine residue is the result of Dicer cleavage heterogeneity or is the result of non-templated nucleotide addition. In the case of miR-10a/b, which contain the TUTase sequence motif and are uridylated *in vitro*, we were nonetheless unable to rank it in our list of highly uridylated miRNAs because of a genomic thymidine residue after the 3' end of the mature miRNA. Still, upon TUTase depletion this terminal nucleotide is dramatically reduced and the corresponding adenylated species is increased. This situation is found for several of our predicted miRNAs suggesting that we may be underestimating the extent of 3' uridylation (and in some cases adenylation) in our studies, yet we cannot rule out that this uridylation reduction is due to a change in Dicer processing.

A recent study described a role for GLD2/PAPD4/TUTase 2 in stabilizing a subset of mature miRNAs in cultured cells (D'Ambrogio et al., 2012). GLD2 was shown previously to stabilize mature miR-122 by adding a single non-templated adenosine residue in mouse liver and mouse embryonic fibroblasts (MEFs), and the recent study extended that work to show stabilization of several other miRNAs by a similar mechanism (D'Ambrogio et al., 2012; Katoh et al., 2009). Interestingly, a subset of miRNAs that are responsive to GLD2 mediated adenylation are specified by sequences in their 3' ends and include some but not all let-7 family members (e.g. let-7i, identified as TUTase-dependent in our study). The authors also found that miR-99a and miR-196a are stabilized by GLD2 despite lacking a 3' GLD2 stabilizing sequence

(D'Ambrogio et al., 2012). The overlap between GLD2-sensitive miRNAs and TUTase-dependent miRNAs suggests that GLD2 may be the PAP responsible for compensatory adenylation in the absence of Zcchc11 and Zcchc6. This adenylation activity may stabilize miRNAs in the absence of a uridylation residue and maintain levels of functional miRNAs. Alternatively, there may be a miRNA sensing mechanism that maintains a constant population of unmodified miRNAs, thus replacing uridylated species with adenylated ones.

Our deep sequencing results largely support our *in vitro* analysis but there are notable exceptions including the general absence of many uridylated mature let-7 family members. While uridylation is detected above background for let-7/e/i, this is complicated by adjacent genomic thymidine residues for most let-7 species including let-7a/b/c/d/f/g. For example, reads of let-7b contain a 3' terminal U roughly 15% of the time, and this decreases 3-fold after TUTase depletion. Correspondingly, 3' adenylation makes up 18% of reads in control cells compared to over 26% in shTUT cells, following the trend of other predicted miRNAs such as miR-99a. Similar issues exist for several other predicted miRNAs and therefore may significantly underestimate the extent of 3' uridylation and/or adenylation. Finally, our analysis focuses primarily on 5p-miRs as all of our predicted substrate miRNAs are fortuitously derived from the 5' arm of miRNA precursors. Many 3p- miRs were found to be significantly uridylated in our sequencing study (data not shown) but separating this mark from pre-miRNA effects is impossible at this time. Therefore there are likely at least a few 3p-miRNAs that are subjected to Zcchc11- and Zcchc6-dependent terminal uridylation. Elucidating the full panoply of modified miRNAs and understanding the consequences of this activity are exciting questions in this growing field and warrant further investigation as an insight to an important form of gene regulation.

MATERIALS AND METHODS

Uridylation assays

In vitro uridylation assays were performed as described previously except with a final RNA concentration of 300nM (Thornton et al., 2012).

RNA Oligo	Sequence
let-7g guide	UGAGGUAGUAGUUUGUACAGUU
let-7g passenger	CUGUACAGGCCACUGCCUUGC
GL2 guide	UCGAAGUAUUCCGCGUACGUU
GL2 passenger	CGUACGCGGAAUACUUCGAUU
let-7i guide	UGAGGUAGUAGUUUGUGCUGUU
let-7i passenger	CUGCGCAAGCUACUGCCUUGCU
let-7i domains 1/2 mut	UAGUCGCUGCAUUUGUGCUGUU
let-7i domains 2/3 mut	UGAGGUAUGCAGCCUAGCUGUU
let-7i guide 1/2/3 mut	UAGUCGCUGCAGCCUAGCUGUU
let-7i delG	UUACAUACUAAUUUCUACUCUU
GL2 with let-7 motif	UCGAAGUGUAGUUUGUACGUU
miR-10a guide	UACCCUGUAGAUCCGAAUUUGUG
miR-26a guide	UUCAAGUAAUCCAGGAUAGGC
miR-10a double mut	UACCCUUGCAAUCCGAAGCCUAG
let-7g+A guide	UGAGGUAGUAGUUUGUACAGUUA
let-7g+U guide	UGAGGUAGUAGUUUGUACAGUUU

Cell culture

P19 cells were cultured in αMEM+10% FBS with daily addition of 100nM retinoic acid over the time course described.

Mouse tissue extraction

Tissues and organs were collected from >5 month-old mice and were homogenized. Lysates were cleared and quantitated using Bradford reagent before immunoblotting.

Lentivirus production and infection

Lentivirus was prepared according to the manufacturer's protocol (Invitrogen #K4975-00) and supernatant was filtered through a 0.45μm filter before being stored at -80C or used immediately. Hela cells were transduced with 500ul of lentivirus supernatant in the presence of polybrene (4μg/ml) and incubated overnight. Media containing either Puromycin (shZcchc11, shGFP, 2.5μg/ml) Hygromycin (shZcchc6, 140μg/ml) or both (shZcchc11/6) was added 36 hours post-infection and stably resistant cell pools were grown before analysis.

shRNAs used in this study

Plasmids	Identifier	Targeted sequence	Resistance
pLKO+shGFP	Sigma SHC005		Puro
pLKO+shZcc11 #1	Sigma #TRCN0000150277	TCAGTTACATTCAGCAGAAA	Puro
pLKO+shZcc11 #2	Sigma #TRCN0000146303	CGTGATAGTGATCTGGATATT	Puro
pLKO+shZcchc6 #408	Custom	GGAATTGCTGCGGTTCTATGC	Hygro
pLKO+shZcchc6 #409	Custom	GTGACCTTGACGTCTGTATGA	Hygro

Small RNA isolation and cloning

Small RNAs were cloned according to the manufacturer's protocol (Epibio #SMMP101212). Briefly, 50ug of Trizol-purified total RNA (Invitrogen) was fractionated using miRvana columns (Invitrogen) and 300ng of <200nt RNA was used in the cloning reactions. Resulting libraries were gel purified from a NativePage 4-16% Bis-tris acrylamide gel (Life Technologies #BN1002BOX) and subjected to Illumina sequencing. 18-30nt small RNA sequences were used for analysis, after removing the adaptor sequences. Bowtie program (http://bowtie-bio.sourceforge.net/index.shtml) was used for aligning small RNAs to annotated human miRNA sequences in miRBase database (http://www.mirbase.org), and two mismatches were permitted. MicroRNA sequences with one "U" and "A" mismatches at the 3' end were considered as untemplated uridylated and adenylated miRNA, respectively.

Expression plasmids

Zcchc11 expression constructs were cloned as described(Thornton et al., 2012).

Primer	Sequence
Zcchc11 PUPΔZn F XhoI	TGCCGCCTCGAGAAAATTGATCTAAAACCTCTACCACCAATG
Zcchc11 PUPΔZn R SalI	TGCCGCGTCGACATGATTCAAGTCAAAAGGATCTTCAATTG

Antibodies

The following antibodies were used for western blots in this study: α-Zcchc11 (Proteintech Group #18980-1-AP), α-Zcchc6 (OpenBioystems custom rabbit polyclonal antibody), α-Flag (Sigma #A8592), α-Lin28A (Cell Signaling #3978), α-Tubulin (Abcam #AB6046), α-Actin (Sigma #A2066).

REFERENCES

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Blahna, M.T., Jones, M.R., Quinton, L.J., Matsuura, K.Y., and Mizgerd, J.P. (2011). Terminal uridyltransferase enzyme zcchc11 promotes cell proliferation independent of its uridyltransferase activity. J Biol Chem *286*, 42381-42389.

Burroughs, A.M., Ando, Y., de Hoon, M.J., Tomaru, Y., Nishibu, T., Ukekawa, R., Funakoshi, T., Kurokawa, T., Suzuki, H., Hayashizaki, Y.*, et al.* (2010). A comprehensive survey of 3' animal miRNA modification events and a possible role for 3' adenylation in modulating miRNA targeting effectiveness. Genome Res *20*, 1398-1410.

Chang, H.M., Triboulet, R., Thornton, J.E., and Gregory, R.I. (2013). A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28-let-7 pathway. Nature *497*, 244-248.

Chiang, H.R., Schoenfeld, L.W., Ruby, J.G., Auyeung, V.C., Spies, N., Baek, D., Johnston, W.K., Russ, C., Luo, S., Babiarz, J.E.*, et al.* (2010). Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. Genes Dev *24*, 992-1009.

D'Ambrogio, A., Gu, W., Udagawa, T., Mello, C.C., and Richter, J.D. (2012). Specific miRNA stabilization by Gld2-catalyzed monoadenylation. Cell Rep *2*, 1537-1545.

Hagan, J.P., Piskounova, E., and Gregory, R.I. (2009). Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nat Struct Mol Biol *16*, 1021-1025.

Heo, I., Ha, M., Lim, J., Yoon, M.J., Park, J.E., Kwon, S.C., Chang, H., and Kim, V.N. (2012). Mono-uridylation of pre-microRNA as a key step in the biogenesis of group II let-7 microRNAs. Cell *151*, 521-532.

Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol Cell *32*, 276-284.

Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell *138*, 696-708.

Hornstein, E., Mansfield, J.H., Yekta, S., Hu, J.K., Harfe, B.D., McManus, M.T., Baskerville, S., Bartel, D.P., and Tabin, C.J. (2005). The microRNA miR-196 acts upstream of Hoxb8 and Shh in limb development. Nature *438*, 671-674.

Jones, M.R., Blahna, M.T., Kozlowski, E., Matsuura, K.Y., Ferrari, J.D., Morris, S.A., Powers, J.T., Daley, G.Q., Quinton, L.J., and Mizgerd, J.P. (2012). Zcchc11 uridylates mature miRNAs to enhance neonatal IGF-1 expression, growth, and survival. PLoS Genet *8*, e1003105.

Jones, M.R., Quinton, L.J., Blahna, M.T., Neilson, J.R., Fu, S., Ivanov, A.R., Wolf, D.A., and Mizgerd, J.P. (2009). Zcchc11-dependent uridylation of microRNA directs cytokine expression. Nat Cell Biol *11*, 1157-1163.

Katoh, T., Sakaguchi, Y., Miyauchi, K., Suzuki, T., Kashiwabara, S., and Baba, T. (2009). Selective stabilization of mammalian microRNAs by 3' adenylation mediated by the cytoplasmic poly(A) polymerase GLD-2. Genes Dev *23*, 433-438.

Kwak, J.E., and Wickens, M. (2007). A family of poly(U) polymerases. RNA *13*, 860-867.

Loughlin, F.E., Gebert, L.F., Towbin, H., Brunschweiger, A., Hall, J., and Allain, F.H. (2011). Structural basis of pre-let-7 miRNA recognition by the zinc knuckles of pluripotency factor Lin28. Nat Struct Mol Biol.

Lunde et al. Crystal structures of the Cid1 poly (U) polymerase reveal the mechanism for UTP selectivity. Nucleic Acids Research (2012) vol. 40 (19) pp. 9815-24

Martin, G., and Keller, W. (2007). RNA-specific ribonucleotidyl transferases. RNA *13*, 1834-1849.

Minoda, Y., Saeki, K., Aki, D., Takaki, H., Sanada, T., Koga, K., Kobayashi, T., Takaesu, G., and Yoshimura, A. (2006). A novel Zinc finger protein, ZCCHC11, interacts with TIFA and modulates TLR signaling. Biochem Biophys Res Commun *344*, 1023-1030.

Nam, Y., Chen, C., Gregory, R.I., Chou, J.J., and Sliz, P. (2011). Molecular Basis for Interaction of let-7 MicroRNAs with Lin28. Cell *147*, 1080-1091.

Newman, M.A., Mani, V., and Hammond, S.M. (2011). Deep sequencing of microRNA precursors reveals extensive 3' end modification. RNA *17*, 1795-1803.

Newman, M.A., Thomson, J.M., and Hammond, S.M. (2008). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. RNA *14*, 1539-1549.

Pasquinelli, A.E., Reinhart, B.J., Slack, F., Martindale, M.Q., Kuroda, M.I., Maller, B., Hayward, D.C., Ball, E.E., Degnan, B., Muller, P.*, et al.* (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. Nature *408*, 86-89.

Piskounova, E., Polytarchou, C., Thornton, J.E., Lapierre, R.J., Pothoulakis, C., Hagan, J.P., Iliopoulos, D., and Gregory, R.I. (2011). Lin28A and Lin28B Inhibit let-7 MicroRNA Biogenesis by Distinct Mechanisms. Cell *147*, 1066-1079.

Rau, F., Freyermuth, F., Fugier, C., Villemin, J.P., Fischer, M.C., Jost, B., Dembele, D., Gourdon, G., Nicole, A., Duboc, D.*, et al.* (2011). Misregulation of miR-1 processing is associated with heart defects in myotonic dystrophy. Nat Struct Mol Biol *18*, 840-845.

Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E.E., Nitsch, R., and Wulczyn, F.G. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. Nat Cell Biol *10*, 987-993.

Saitoh, S., Chabes, A., McDonald, W.H., Thelander, L., Yates, J.R., and Russell, P. (2002). Cid13 is a cytoplasmic poly(A) polymerase that regulates ribonucleotide reductase mRNA. Cell *109*, 563-573.

Schmidt, M.J., West, S., and Norbury, C.J. (2011). The human cytoplasmic RNA terminal U-transferase ZCCHC11 targets histone mRNAs for degradation. RNA *17*, 39-44.

Siomi, H., and Siomi, M.C. (2010). Posttranscriptional regulation of microRNA biogenesis in animals. Mol Cell *38*, 323-332.

Thornton, J.E., Chang, H.M., Piskounova, E., and Gregory, R.I. (2012). Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). RNA *18*, 1875-1885.

Thornton, J.E., and Gregory, R.I. (2012). How does Lin28 let-7 control development and disease? Trends Cell Biol *22*, 474-482.

Trippe, R., Guschina, E., Hossbach, M., Urlaub, H., Luhrmann, R., and Benecke, B.J. (2006). Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. RNA *12*, 1494-1504.

Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. Science *320*, 97-100.

Wee, L.M., Flores-Jasso, C.F., Salomon, W.E., and Zamore, P.D. (2012). Argonaute divides its RNA guide into domains with distinct functions and RNA-binding properties. Cell *151*, 1055-1067.

Wyman, S.K., Knouf, E.C., Parkin, R.K., Fritz, B.R., Lin, D.W., Dennis, L.M., Krouse, M.A., Webster, P.J., and Tewari, M. (2011). Post-transcriptional generation of miRNA variants by multiple nucleotidyl transferases contributes to miRNA transcriptome complexity. Genome Res *21*, 1450-1461.

Yates, L.A., Norbury, C.J., and Gilbert, R.J. (2013). The long and short of microRNA. Cell *153*, 516-519.

Yekta, S., Tabin, C.J., and Bartel, D.P. (2008). MicroRNAs in the Hox network: an apparent link to posterior prevalence. Nat Rev Genet *9*, 789-796.

CHAPTER 4

Discussion, Unpublished Data, Conclusions

The author produced all text and figures in this chapter. As of the publication of this thesis none of these data have been published.

DISCUSSION

Implications of TUTase domain analysis: structure-function relationships

The work described in this thesis covers the identification and characterization of a novel class of nucleotide transferases, including their functional domain architecture and the molecular mechanism of their substrate specification. This work has expanded upon our laboratory's central theme of miRNA regulation in embryonic stem cells—and specifically the role of Lin28 proteins in let-7 biogenesis—to include how mechanistically these enzymes recognize their targets, and that TUTases are endowed with a sequence-specific nucleotide recognition capability. This work developed from a biochemical perspective using *in vitro* methods and structure-function relationships to draw a line between RNA metabolism, stem cell biology, vertebrate development, and cancer. The route from project inception to conclusions was tortuous but by connecting these disparate fields this work shows the power of biochemical analysis in revealing the complexity of biological systems. Tremendous work remains to be done in uncovering the precise biochemical mechanism underlying Lin28 and let-7 regulation on the one hand, and the physiological impact of mature miRNA uridylation on the other. This work should form a foundation upon which other researchers from our lab and other groups can continue to contribute to the field of gene silencing.

The early stages of this work formed around one simple question: How does the multidomain TUTase Zcchc11 function to repress let-7 miRNAs in embryonic stem cells? Given the sole role ascribed to Zcchc11 at that point it seemed evolutionarily profligate to allow for the conservation of such a large protein when several related enzymes were known to perform similar functions while being much smaller overall. The unique architecture of Zcchc11 (and soon, we learned, Zcchc6) suggested that these enzymes might have other functions necessitating

such complexity. In parallel several groups have uncovered such divergent functions, namely the discovery that Zcchc11 can alter the cell cycle in cancer cells independent of its uridyl transferase activity (Blahna et al., 2011). Overexpression of the N-terminal region of Zcchc11 that has no enzymatic activity was sufficient to disrupt the cell cycle to the same extent as overexpressing the full-length protein. Similarly, in the study outlining the initial identification of Zcchc11, a short cDNA was cloned in macrophages that corresponded to an N-terminal region similar to that used in the study by Blahna and colleagues (Minoda et al., 2006; Blahna et al., 2011). From experiments described in this thesis we found this N-terminal region was indeed insufficient to support enzymatic activity, but that it was absolutely required to support robust uridylation activity *in vitro*. This N-terminal region, which contains a unique C2H2 zinc finger as well as a TRF4 domain highly similar to the defined active site, is also essential for Lin28-mediated uridylation of let-7 miRNA precursors. Aside from the necessity of the TRF4 domain in supporting basal activity, the C2H2 zinc finger is sufficient to explain the functional interaction between Zcchc11 and Lin28 (Figure 2.1). The regional conservation of domains that mediate cell cycle changes, support basal uridylation activity, and promote binding partner-specific activity—all while being distal from the active site itself—underscores the importance of the full-length protein in carrying out at least some of its cellular functions. Given that this distally conserved region is highly homologous to the verified active site also suggests that it may be involved in substrate binding or positioning, and perhaps arose as a result of an intramolecular gene duplication event. Of the mammalian non-canonical poly(A) polymerases, only Zcchc6 and Zcchc11 have this unique domain arrangement, indicating the importance of having redundant enzymes that carry out specialized roles.

A recent study of the *Xenopus laevis* ortholog of Zcchc6 (XTUT7) supported many of the

findings described in Chapter 2, but also expanded upon our conclusions. Lapointe et al. (2013) used RNA-protein tethering assays to monitor XTUT7 activity *in vitro* and found many of the regions both distal and proximal to the active site to be dispensable for robust uridylation activity. Interestingly they found the TRF4 domain to have no effect on XTUT7 activity in their hands, despite a high level of overall similarity with human Zcchc11/TUT4. The authors rightly note, however, that tethering the enzyme to RNA dramatically increases the sensitivity of the assay, so much so in fact that deleting all three CCHC zinc fingers did not completely abolish detectable activity. In this study the authors also found a short basic patch of amino acids near the active site that is sufficient for RNA binding, and upon deletion, leads to the loss of XTUT7 activity in their tethered RNA assay. By using molecular modeling simulations the authors also identified a conserved histidine residue that is positioned near a bound UTP molecule, similar to the precisely placed histidine residue that is essential for UTP selectivity in the *S. pombe* TUTase Cid1 (Yates et al., 2010). Indeed, this residue in XTUT7 (His1269) is essential for nucleotide selectivity, as replacing it with a leucine significantly reduces nucleotide preference *in vitro*. This residue is also essential for XTUT7-mediated translational repression of a tethered transgene, although the relevance of this translation repression *in vivo* is poorly understood. The findings described by Lapointe et al. (2013) firstly underscore the high evolutionary conservation of TUTases across diverse clades. Key residues, whether the active site aspartate triad or the UTP-binding histidine residue, are highly conserved from fission yeast to humans. Furthermore, the complexity and apparent dispensability seen in the domain arrangement of TUTases suggests that these proteins play roles in various pathways. One of the outstanding questions that still remains is how and when these different domains are employed and to what end the cell might use them depending on context. Further investigation is required to answer this complex and

incompletely-answered question.

In examining the domain architecture of Lin28-mediated uridylation we have come to appreciate that it is not conserved in the context of mature miRNA uridylation. While the PUP domain has very low activity in uridylating pre-let-7, the overall preference of Zcchc11 for miRNAs containing GUAG/UUGU is so high that the PUP mutant still carries out detectable uridylation activity, albeit lower than that for the wild-type protein (Figure 3.4C, data not shown). When the three CCHC zinc fingers flanking the active site of PUP are deleted, however, all activity is lost. This result suggests that sequence-dependent uridylation is mediated through structures proximal to the active site, as opposed to the N-terminal regions involved in pre-miRNA uridylation. While this work has helped to dissect the domains required for Zcchc6 and Zcchc11 activity, it remains to be seen how the manifold domains of these TUTases control other realms of TUTase activity.

Implications of TUTase domain analysis: inhibitor screening and structural biology

The importance of understanding various domain contributions of Zcchc11 is evident in addressing the basic question of how these novel enzymes function, but this work has clear applied biological implications as well. As described in Piskounova et al. (2011), depleting Zcchc11 in mouse models of Lin28A-driven cancers is sufficient to cause tumor regression and elevate let-7 levels to an extent similar to knockdown of Lin28A. While the field is actively pursuing small molecule inhibitors against Lin28 proteins, these short RNA-binding proteins are far from ideal target candidates.

Generally, inhibitory molecules target enzymes with active sites of a known general architecture using compounds analogous to endogenous substrates. Envisioning a way to

synthesize a synthetic RNA molecule that is cell permeable, specific to Lin28, and able to potently block both Lin28A and Lin28B underscores the difficulty of targeting these proteins. Zcchc6 and Zcchc11, meanwhile, are two of a very select field of enzymes that exclusively use uridine triphosphate as a substrate, and are closely related to a group of enzymes for which there are published structures (Yates et al., 2012; Deng et al., 2005; Stagno et al., 2010). Furthermore, the domain analysis described here can be used to design a shorter polypeptide to aid in protein purification and crystallization. To this end we have begun collaborations with structural biologists and synthetic chemists as a first step to search for inhibitors of TUTases. We have also designed a high-throughput screening approach using recombinant protein and a luciferase-based detection system to monitor TUTase activity (data not shown). This work, while far from yielding a useful crystal structure, let alone an inhibitor molecule, nonetheless lays the groundwork for further progress on these topics.

Sequence-specific miRNA uridylation as a novel paradigm of gene regulation

The work described in Chapter 3 highlights a novel role for TUTases in targeting specific RNA molecules based entirely on their primary nucleotide sequence. Until the findings described here there was nothing known about TUTase activity controlled by RNA *cis* regulatory elements. While the sequence motif of GUAG/UUGU is both necessary and sufficient for miRNA targeting by Zcchc6 and Zcchc11, and depletion of these TUTases leads to reduced non-templated uridylation for these miRNAs, we have not ruled out that these TUTases target miRNAs containing other sequences. It would be possible to analyze the most-altered U-containing miRNAs after TUTase depletion in various cellular contexts to determine other stimulatory sequences but that is beyond the scope of this study. Furthermore it is possible that

other TUTases or PAPs use distinct sequences to target their substrate miRNAs. If indeed this speculation results in the discovery of other miRNA sequence-specific TUTases, it would bolster our findings and further promote the idea that miRNAs are under evolutionary selection to contain sequences not just required for seed/mRNA binding, but for TUTase targeting as well. To this end it would be of great interest to perform genome-wide analyses to examine TUTase-interacting miRNAs.

In our analysis of sequence-dependent miRNA modification we were capable of predicting only half of the targeted miRNAs *a priori*. Several other 5p-miRs also undergo a reduction in non-templated uridylation, including miR-26a/b as reported previously (Jones et al., 2009). The identification of this miRNA family as a positive control and the understanding that other unrelated miRNAs also have TUTase-dependent modifications suggests that there are other specificity factors—either within the miRNAs themselves or mediated through RNA-binding proteins—that we failed to identify. It is intriguing to note that all of the miRNAs expressed from clusters that contain a member of the let-7 family, the genomically-linked miR-125 family, or the miR-99/100 family undergo TUTase-dependent reductions in uridylation despite the miR-125 family lacking the predicted target sequence (Figures 3.3B & 3.7D). Connecting locus-based identification to mature miRNA modification would require a difficult-to-imagine mechanism that "remembers" where in the genome an individual miRNA was encoded and "tracks" its mature product so that it is modified by a TUTase. Instead, we favor the idea that there are as-yet unidentified factors involved in specifying miRNAs for non-templated uridylation.

TUTases as developmental regulators: Hox genes and vertebrate patterning

Our analysis of mature miRNA uridylation by Zcchc6 and Zcchc11 led us to the

surprising realization that a select subset of miRNAs contains a sequence motif required for uridylation targeting. Even more surprising was the finding that these miRNAs are primarily known as developmental regulators and often target Hox genes in a temporally and spatially controlled manner. Hox genes are well-established specificiers of animal patterning, and perturbations in Hox gene expression can grossly alter the regional developmental patterns along the anterior-posterior axis (Nüsslein-Volhard & Wieschaus, 1980). The Hox-targeting miRNAs identified in our study have been shown to preferentially repress anterior patterning genes in favor of posterior gene expression. This "posterior prevalence" as described in Yekta et al. (2008) suggests that Hox-targeting miRNAs reinforce the evolutionarily dominance of posterior patterning in bilaterian development. If uridylation attenuates miRNA function as described previously (Jones et al., 2009; Jones et al., 2012), we predict that TUTase depletion in a developing organism will enhance miRNA activity against their Hox gene targets and disrupt posterior development. To test this we have begun a collaboration with the laboratory of Leonard Zon to understand the role of TUTases in zebrafish development. This powerful model organism permits medium-throughput screening to monitor changes in vertebrate development, and because its developmental stages are well known and easily synchronized, we can observe hundreds of embryos per experiment and perform quantitative experiments including RT-qPCR and deep sequencing of small RNAs. We are in the early stages of these experiments but have designed and procured morpholinos targeting the zebrafish orthologs of Zcchc6 and Zcchc11. These experiments should expand upon our cell-based examination of developmental miRNAs, and provide a dynamic picture of the role of TUTases *in vivo*.

SELECTED UNPUBLISHED DATA

Zcchc6 and Zcchc11 are exclusively cytoplasmic

Our findings that the redundant TUTases Zcchc11 and Zcchc6 function together to mediate the degradation of let-7 precursors in embryonic stem cells suggests their activity is localized at least partially to the cytoplasm where miRNA precursors are retained. Zcchc11 has been implicated in both nuclear and cytoplasmic pathways apart from pre-let-7 uridylation (Minoda et al., 2006; Jones et al., 2009) but we wished to examine how these enzymes are localized in a human cancer cell line. Hela cells were fractionated and Western blotting for the endogenous proteins revealed exclusive cytoplasmic localization for both Zcchc6 and Zcchc11 (Figure 4.1). This supports our findings of mature miRNA uridylation as described in Chapter 3. The sole report of nuclear Zcchc11 localization was in a macrophage cell line after stimulation with lipopolysaccharide (LPS) (Minoda et al., 2006). It remains possible that these TUTases have other roles in specific cellular contexts or that their localization and/or function may be dynamically regulated.

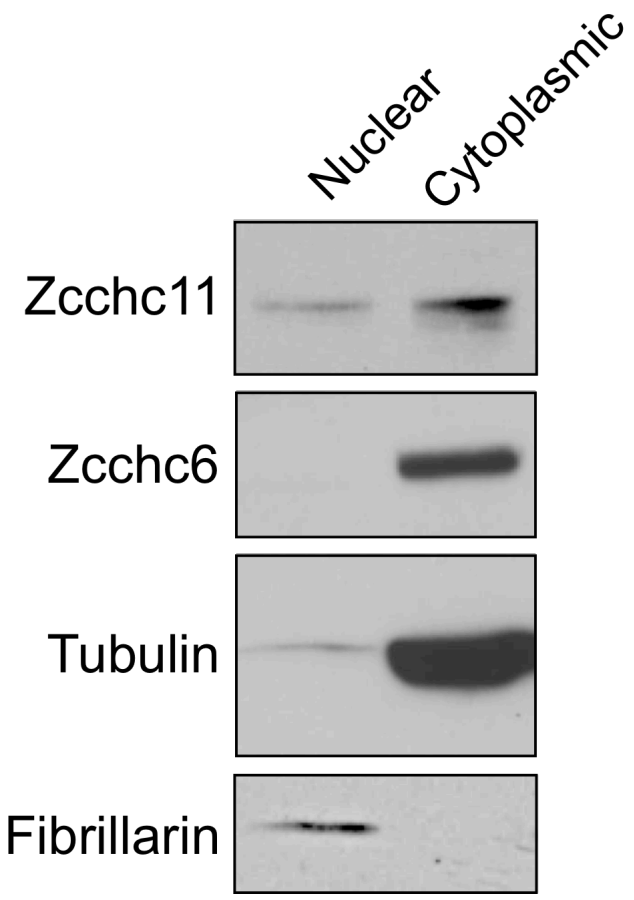


Figure 4.1: Compartmentalization of Zcchc6 and Zcch11

Hela cells were fractionated and analyzed by Western blot. Tubulin is a marker of the cytoplasmic fraction and fibirillarin marks the nucleus.

Zcchc11 resides in a high molecular weight complex in a human cancer cell line

In the early days of investigating Zcchc11 activity in contexts other than embryonic stem cell regulation, we screened a panel of cell lines for their expression of Zcch11 and Lin28 protein family members. Focusing on the human non-small cell lung cancer (NSCLC) cell line H1299, which expresses high levels of Zcchc11 (See Figure 3.6A), we created a stable line that expressed Flag-tagged human Zcchc11. We cultured 60 15cm plates of this stable cell line, prepared whole cell extracts, and performed a large-scale Flag immunoprecipitation. Eluted fractions were examined by silver stain and Western blot using antibodies against Zcchc11 and Flag epitopes (Figure 4.2A). Signal peaked in elution 2, and this fraction was analyzed by tandem Mass Spectrometry (MS-MS).

To determine the approximate molecular weight of Zcchc11-containing complexes, the remainder of Elution 2 was concentrated and purified on a Superose 6 size exclusion column using FPLC (Figure 4.2A). The Flag-tagged Zcchc11 protein migrates as a ~185kDa monomer under denaturing conditions, but elutes in a multiprotein complex peaking at ~700kDa under non-denaturing but stringent elution conditions of 500mM KCl, suggesting it stably interacts with numerous protein and/or RNA species. Although the list of Zcchc11 co-immunoprecipitating proteins is extensive, this experiment primarily sought to identify a uridylation-specific nuclease, which has since been identified as Dis3l2. Although this approach did not lead to a viable candidate, this list of interacting proteins may become useful in investigating protein factors that mediate let-7-independent roles of Zcchc11.

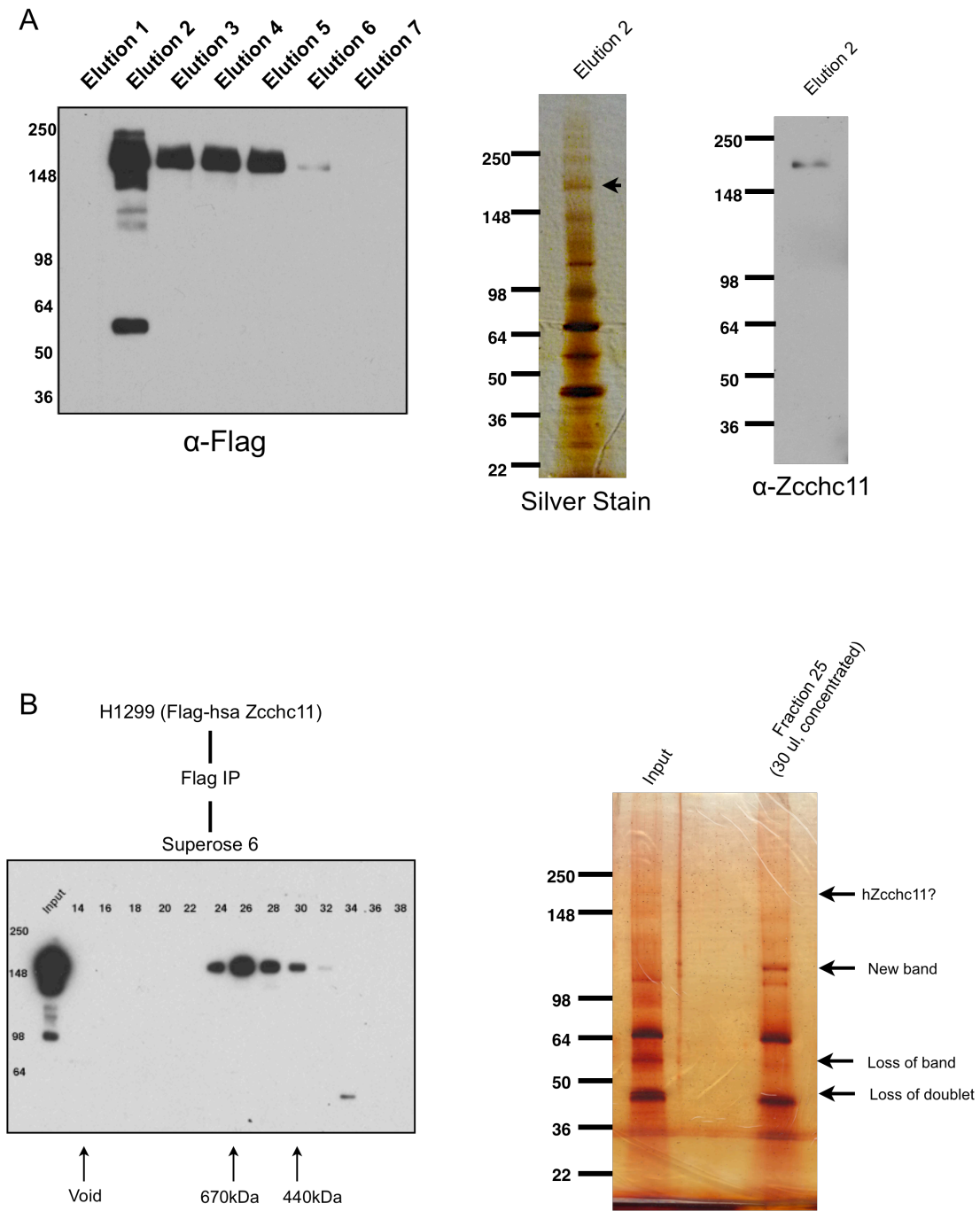


Figure 4.2: Flag Zcchc11 resides in a large molecular weight complex

Figure 4.2 continued

A. H1299 human cancer cells were stably transfected with human Zcchc11 and large-scale Flag immunoprecipitates were purified. Purified protein peaks in Elution 2 (left), and Zcchc11 is detectable by silver stain and α-Zcchc11 Western blot (right). Elution 2 was later used for tandem mass spectrometry analysis.

B. Flag-Zcchc11 complex was purified on a Superose 6 size-exclusion column and resulting fractions were analyzed by Flag Western blot (left). Concentrated Fraction 25 was analyzed by silver stain, which showed a differential staining pattern compared to the input fraction.

Developmental regulation of Zcchc11 is not mediated through let-7 miRNAs

As shown in Figure 3.5, Zcchc11 is potently down-regulated during retinoic acid (RA)-mediated differentiation *in vitro*. During this differentiation protocol Lin28 levels decrease (Figure 3.5B) and let-7 levels increase (Figure 4.3A), as reported previously (Balzer et al., 2010, Heo et al., 2009; Viswanathan et al., 2009). Intriguingly, outdated versions of the miRNA target prediction algorithm Targetscan show a putative let-7 binding site in the 3'UTR of Zcchc11. To test whether Zcchc11 is regulated by let-7 miRNAs during development, we cloned the entire mouse Zcchc11 3'UTR in a luciferase reporter construct and tested whether it responds to changes in let-7 levels.

Hela cells, which express high levels of let-7 miRNAs, were transfected with either an empty vector reporter; a reporter containing the 3'UTR of Zcchc11; or a reporter with three perfect let-7 sites, which is potently repressed in the presence of let-7 miRNAs. These cells were co-transfected with either control antisense inhibitory RNA (antagomir) or an antagomir targeting let-7. As shown in Figure 4.3B the reporter containing let-7 sites was dramatically derepressed after let-7 inhibition, while the empty vector and the construct containing the Zcchc11 3'UTR were unchanged after let-7 depletion, indicating that the murine Zcchc11 3'UTR is not regulated by this family of miRNAs. It remains unknown how Zcchc11 is down-regulated during development but it may be a transient transcriptional decrease that is selectively relieved in adult tissues destined to express Zcchc11, or it may be through post-translation modifications as has been suggested in at least one study (Minoda et al., 2006).

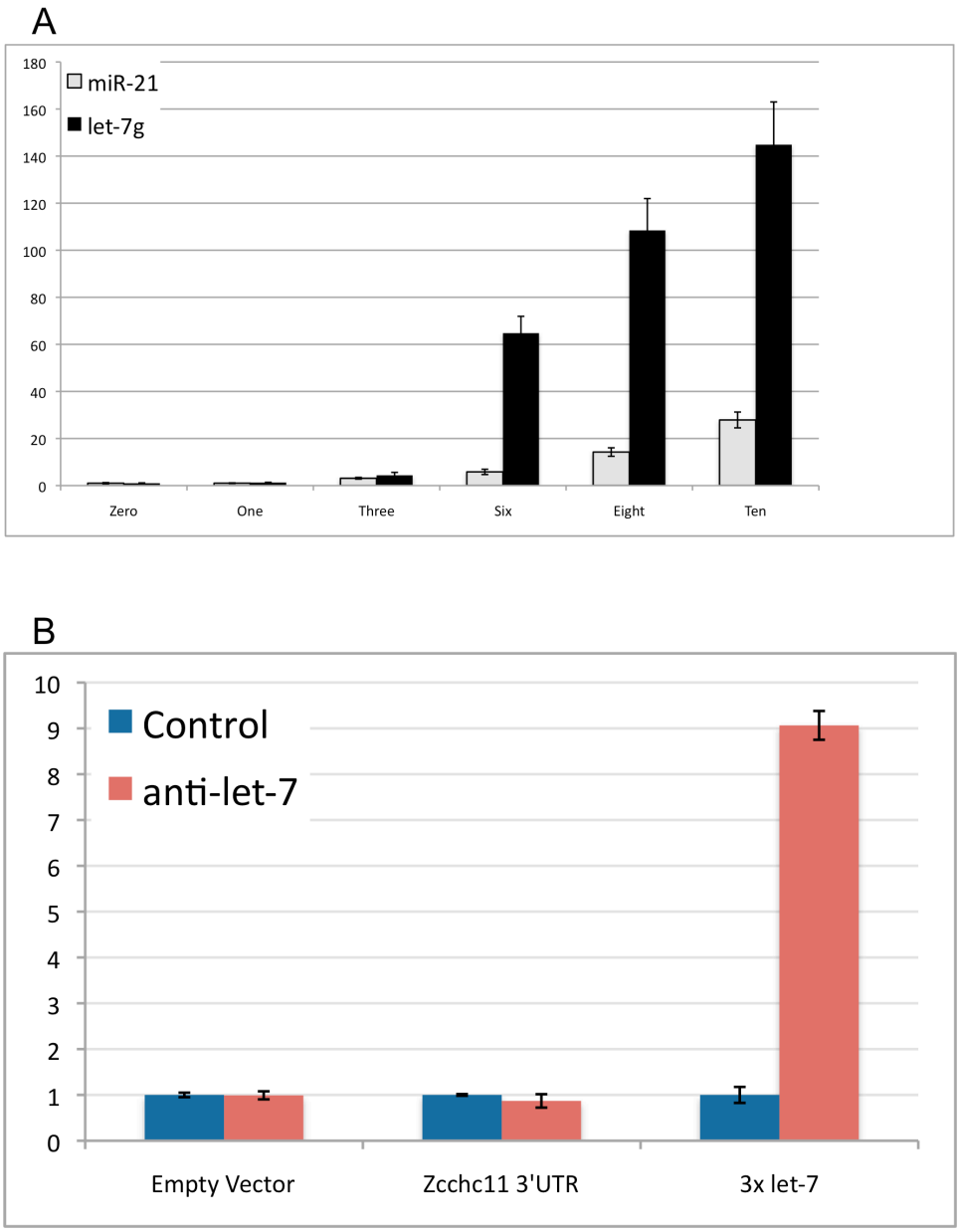


Figure 4.3: Zcch11 is not regulated by let-7 miRNAs

A. After RA-mediated differentiation, let-7 levels are dramatically increased.

B. Luciferase assay showing let-7-independent regulation of the murine Zcchc11 3'UTR. A reporter with 3 perfect let-7 binding sites was used as a positive control.

Lin28-mediated uridylation of pre-let-7 requires both a stem and a loop structure

While investigating the structural requirements of RNA substrates for Lin28 and Zcchc11-mediated uridylation, we began by determining whether either the stem region or the loop region was sufficient for this activity. It was known that Lin28 is a specificity factor that recognizes sequences within the let-7 preE, but it was unknown whether Zcchc11 conferred any extra specificity for this interaction. To determine if this was the case, we performed additional uridylation assays as in Figure 2.1 using pre-let-7g, let-7g preE, and let-7g duplex as substrates. As shown in Figure 4.4, pre-let-7g was sufficient to mediate Lin28-enchanced uridylation, but neither the stem nor the loop was capable of eliciting this activity. This finding suggests that Zcchc11 has no obvious sequence preference for different dsRNA components of the let-7 pre-miRNA, and that Lin28 binding to a miRNA's preE is sufficient to drive enhanced TUTase activity (see also Figure 2.2). While this result did not elucidate mechanistic information on Zcch11 function, it did lead us to test whether either the guide or passenger strands were preferred substrates of these TUTases, and led to our discovery of sequence-specific uridylation of single-stranded mature miRNAs (see Chapter 3, especially Figure 3.1).

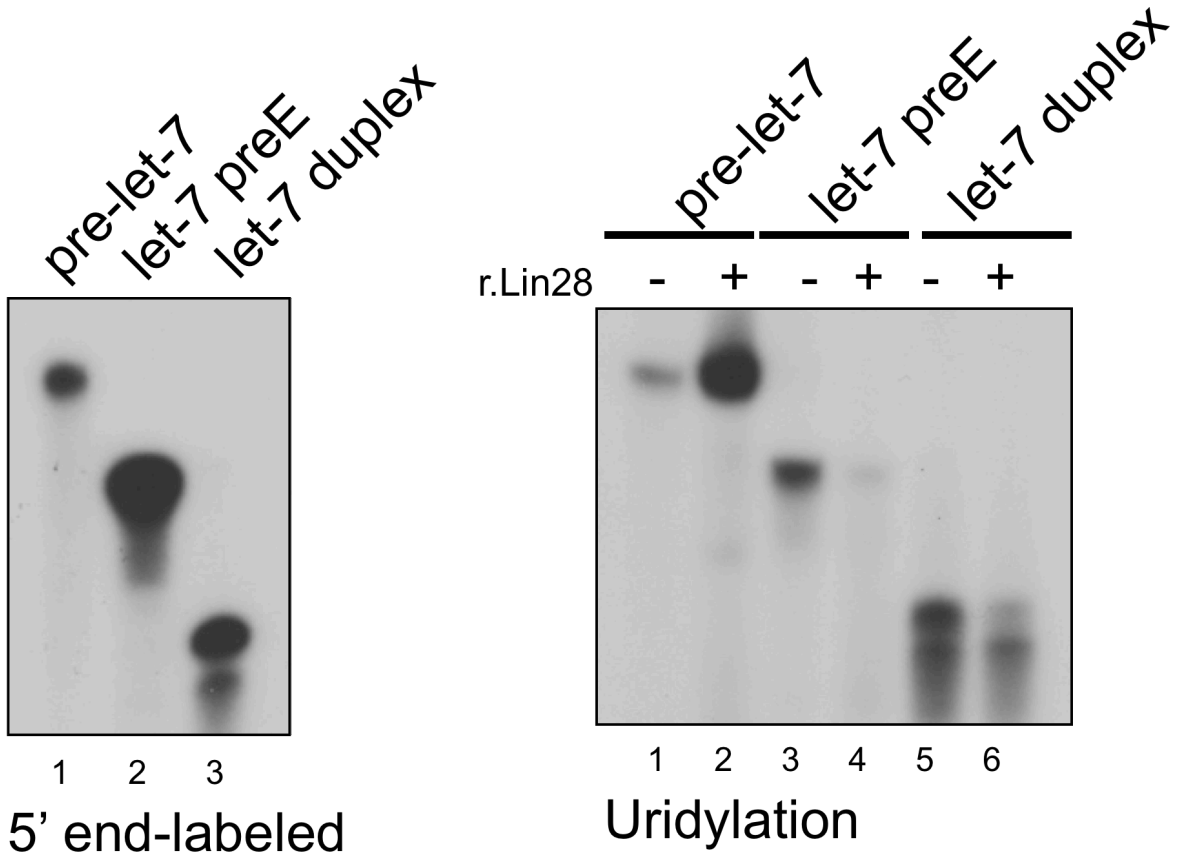


Figure 4.4: Lin28 requires both the stem and loop of let-7 miRNAs to drive Zcchc11-mediated uridylation

CONCLUSIONS

With recent mechanistic discoveries underlying the ancient Lin28/let-7 pathway, the field of stem cell biology was introduced to the core importance of TUTases in regulating gene expression. The work from our lab and several others around the world has expanded on this work to show the importance of TUTases in extremely diverse fields of biology. In the last four years alone, the relevance of TUTases has grown to include the first high-resolution structures, the first knockout mouse, the importance of RNA uridylation in inflammatory diseases and cancer, and as described here the centrality of substrate identification by Zcchc6 and Zcchc11 in regulating distinct phases of developmental miRNA biogenesis. This work, combined with the last several years' progress on dissecting the mechanism of let-7 temporal regulation, has made gains in one of our laboratory's ultimate goals of finding a small molecule inhibitor targeting the Lin28/let-7 axis. Our group has also remained focused on discovering novel regulatory pathways of non-coding RNAs, and the sequence-dependent uridylation by Zcchc6 and Zcchc11 has opened a novel paradigm in miRNA targeting and regulation. Many outstanding questions remain, but the research described in this thesis—both basic and applied—underscores the central role of miRNA regulation in development and disease.

REFERENCES

Balzer, E., Heine, C., Jiang, Q., Lee, V.M., and Moss, E.G. (2010). LIN28 alters cell fate succession and acts independently of the let-7 microRNA during neurogliogenesis in vitro. Development *137*, 891–900.

Blahna, M.T., Jones, M.R., Quinton, L.J., Matsuura, K.Y., and Mizgerd, J.P. (2011). Terminal uridyltransferase enzyme zcchc11 promotes cell proliferation independent of its uridyltransferase activity. J Biol Chem *286*, 42381–42389.

Deng, J., Ernst, N.L., Turley, S., Stuart, K.D., and Hol, W.G.J. (2005). Structural basis for UTP specificity of RNA editing TUTases from Trypanosoma brucei. Embo J *24*, 4007–4017.

Heo, I., Joo, C., Kim, Y.-K., Ha, M., Yoon, M.-J., Cho, J., Yeom, K.-H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell *138*, 696–708.

Jones, M.R., Blahna, M.T., Kozlowski, E., Matsuura, K.Y., Ferrari, J.D., Morris, S.A., Powers, J.T., Daley, G.Q., Quinton, L.J., and Mizgerd, J.P. (2012). Zcchc11 uridylates mature miRNAs to enhance neonatal IGF-1 expression, growth, and survival. PLoS Genet *8*, e1003105.

Jones, M.R., Quinton, L., Blahna, M.T., Neilson, J., Fu, S., Ivanov, A.R., Wolf, D., and Mizgerd, J. (2009). Zcchc11-dependent uridylation of microRNA directs cytokine expression. Nat Cell Biol.

Minoda, Y., Saeki, K., Aki, D., Takaki, H., Sanada, T., Koga, K., Kobayashi, T., Takaesu, G., and Yoshimura, A. (2006). A novel Zinc finger protein, ZCCHC11, interacts with TIFA and modulates TLR signaling. Biochem Biophys Res Commun *344*, 1023–1030.

Nüsslein-Volhard, C., and Wieschaus, E. (1980). Mutations affecting segment number and polarity in Drosophila. Nature *287*, 795–801.

Piskounova, E., Polytarchou, C., Thornton, J.E., LaPierre, R.J., Pothoulakis, C., Hagan, J.P., Iliopoulos, D., and Gregory, R.I. (2011). Lin28A and Lin28B Inhibit let-7 MicroRNA Biogenesis by Distinct Mechanisms. Cell *147*, 1066–1079.

Stagno, J., Aphasizheva, I., Bruystens, J., Luecke, H., and Aphasizhev, R. (2010). Structure of the mitochondrial editosome-like complex associated TUTase 1 reveals divergent mechanisms of UTP selection and domain organization. J Mol Biol *399*, 464–475.

Viswanathan, S., Powers, J., Einhorn, W., Hoshida, Y., Ng, T., Toffanin, S., O'Sullivan, M., Lu, J., Phillips, L., Lockhart, V., et al. (2009). Lin28 promotes transformation and is associated with advanced human malignancies. Nat Genet.

Yates, L.A., Fleurdépine, S., Rissland, O.S., De Colibus, L., Harlos, K., Norbury, C.J., and Gilbert, R.J.C. (2012). Structural basis for the activity of a cytoplasmic RNA terminal uridylyl transferase. Nature Structural & Molecular Biology *19*, 782–787.

Yekta, S., Tabin, C.J., and Bartel, D.P. (2008). MicroRNAs in the Hox network: an apparent link to posterior prevalence. Nat Rev Genet *9*, 789–796.

Appendix I

Supplementary Figures for Chapter 3

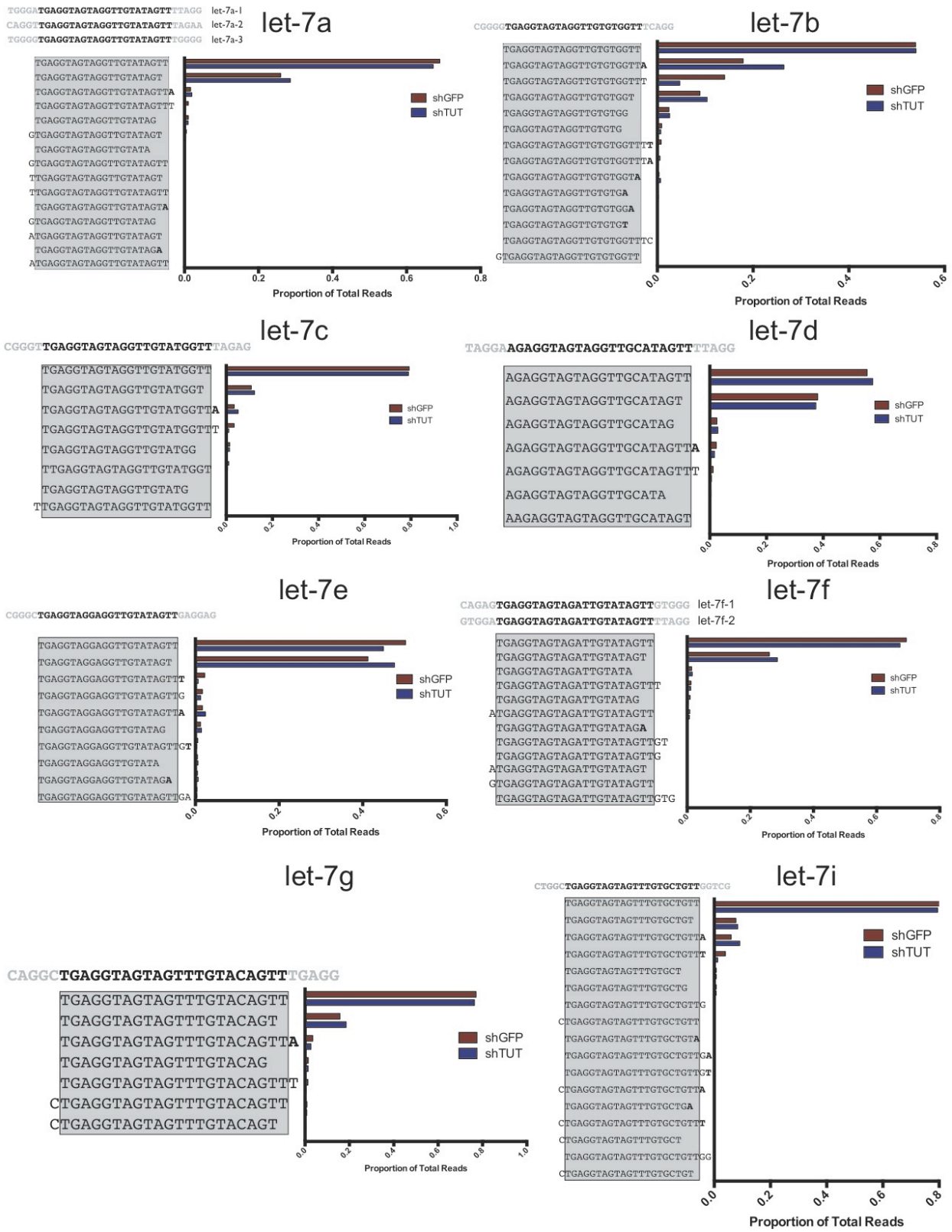
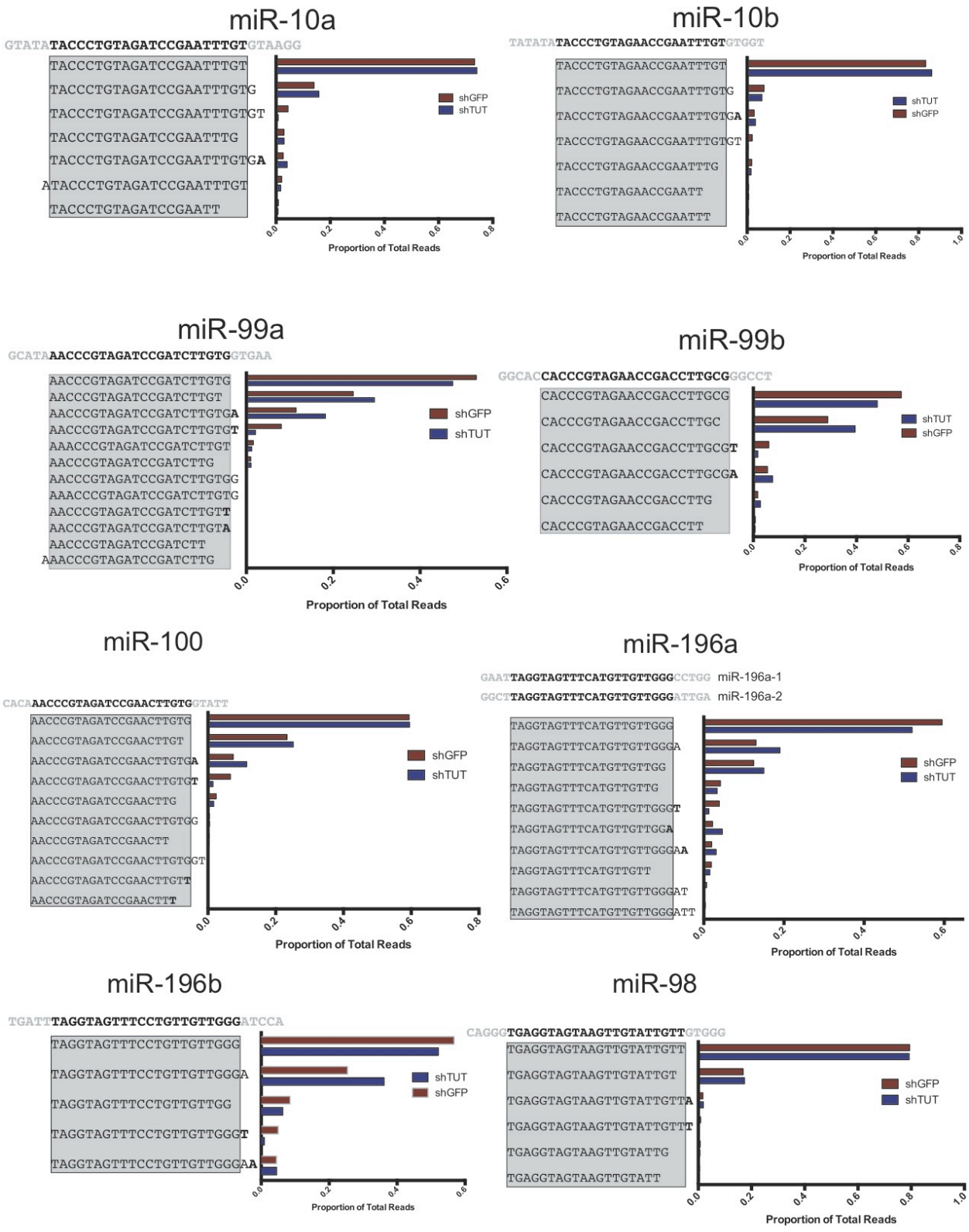


Figure S.1: Length of 3' tails for predicted TUTase-targeted miRNAs

Figure S.1 continued



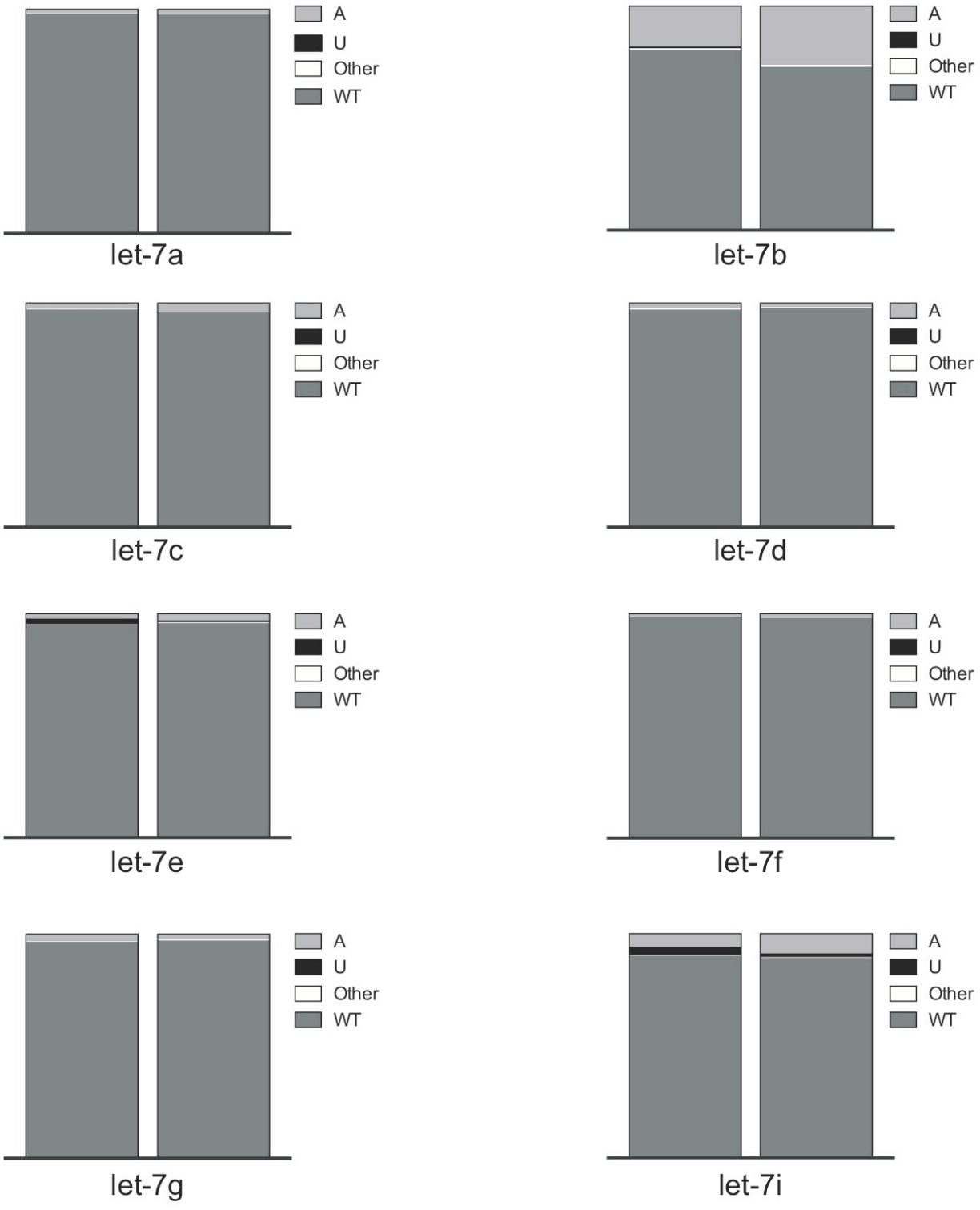
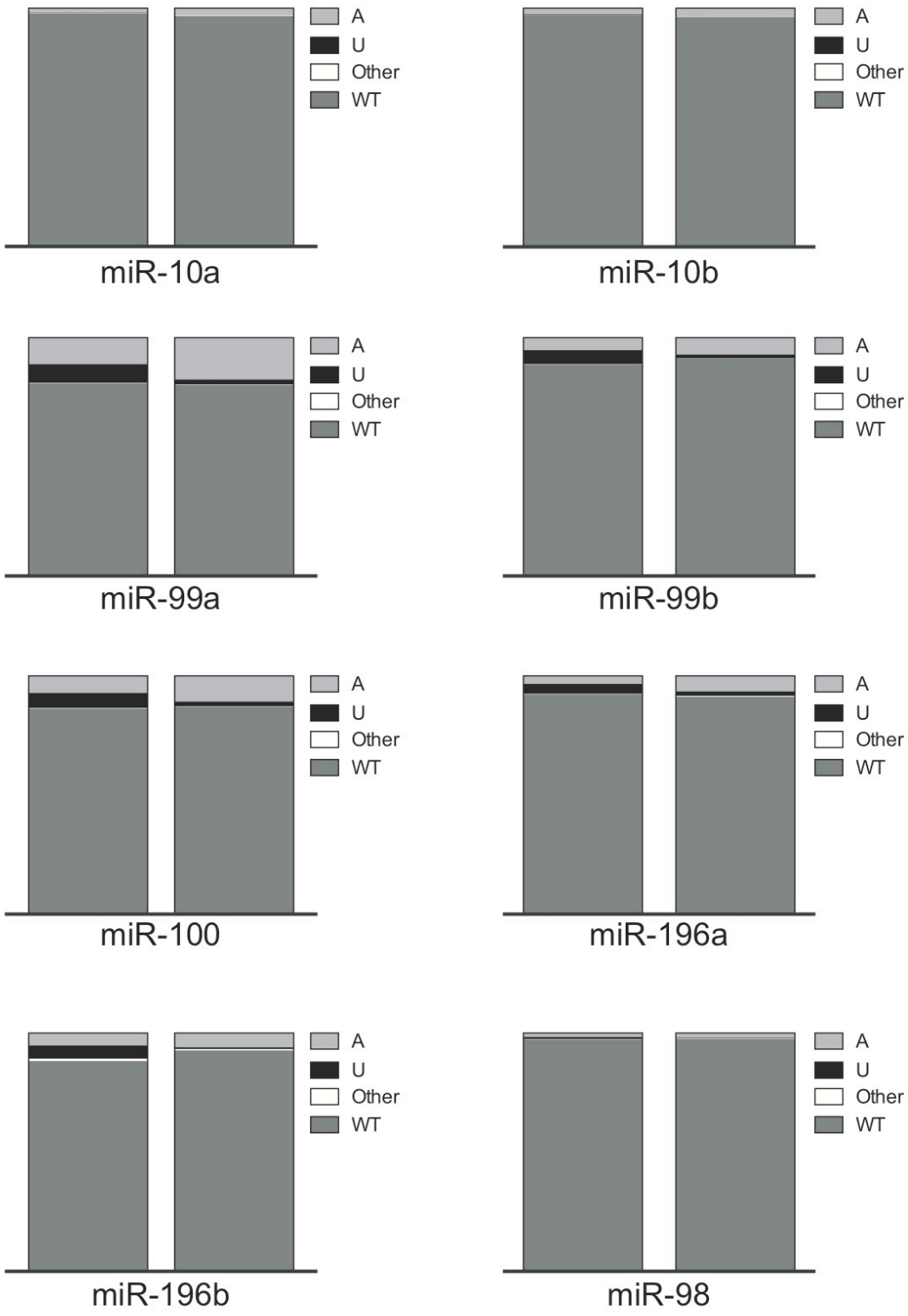


Figure S.2: Composition of terminal nucleotides for predicted TUTase-targeted miRNAs

Figure S.2 continued



Appendix II

Additional Selected Publications

Cell

Lin28A and Lin28B Inhibit let-7 MicroRNA Biogenesis by Distinct Mechanisms

Elena Piskounova,[1,2,4] Christos Polytarchou,[3,5] James E. Thornton,[1,2,4] Robert J. LaPierre,[1,2,4] Charalabos Pothoulakis,[6] John P. Hagan,[1,2,4,7] Dimitrios Iliopoulos,[3,5,*] and Richard I. Gregory[1,2,4,*]
[1]Stem Cell Program, Children's Hospital, Boston, MA 02115, USA
[2]Department of Biological Chemistry and Molecular Pharmacology
[3]Department of Microbiology and Immunobiology
Harvard Medical School, Boston, MA 02115, USA
[4]Harvard Stem Cell Institute, Boston, MA 02115, USA
[5]Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, Boston, MA 02115, USA
[6]Inflammatory Bowel Disease Center, Division of Digestive Diseases, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095, USA
[7]Department of Molecular Virology, Immunology, and Medical Genetics, The Ohio State University Medical Center, Columbus, OH 43210, USA
*Correspondence: dimitrios_iliopoulos@dfci.harvard.edu (D.I.), rgregory@enders.tch.harvard.edu (R.I.G.)
DOI 10.1016/j.cell.2011.10.039

SUMMARY

Lin28A and Lin28B selectively block the expression of let-7 microRNAs and function as oncogenes in a variety of human cancers. Lin28A recruits a TUTase (Zcchc11/TUT4) to let-7 precursors to block processing by Dicer in the cell cytoplasm. Here we find that unlike Lin28A, Lin28B represses let-7 processing through a Zcchc11-independent mechanism. Lin28B functions in the nucleus by sequestering primary let-7 transcripts and inhibiting their processing by the Microprocessor. The inhibitory effects of Zcchc11 depletion on the tumorigenic capacity and metastatic potential of human cancer cells and xenografts are restricted to Lin28A-expressing tumors. Furthermore, the majority of human colon and breast tumors analyzed exclusively express either Lin28A or Lin28B. Lin28A is expressed in HER2-overexpressing breast tumors, whereas Lin28B expression characterizes triple-negative breast tumors. Overall our results illuminate the distinct mechanisms by which Lin28A and Lin28B function and have implications for the development of new strategies for cancer therapy.

INTRODUCTION

Control of gene expression by microRNAs (miRNAs) is important for normal development. Altered miRNA expression is linked with various diseases including cancer (Small and Olson, 2011). miRNA biogenesis begins with transcription of primary transcripts (pri-miRNAs) that contain a stem-loop structure. In the cell nucleus, pri-miRNAs are processed by the Microprocessor, containing the ribonuclease Drosha and its essential cofactor DGCR8 (Denli et al., 2004; Gregory et al., 2004). The Microprocessor cleaves the double-stranded RNA toward the base of the stem loop to release a 60–80 nucleotide (nt) precursor (pre-miRNA) that is exported to the cell cytoplasm and cleaved by Dicer to generate a 22 nt duplex (Hutvágner et al., 2001; Krol et al., 2010). One RNA strand is bound by Argonaute and incorporated into the RNA-induced silencing complex (RISC) (Gregory et al., 2005; Liu et al., 2004). Basepairing between the miRNA and target mRNA guides RISC to complementary transcripts, leading to gene repression through mRNA degradation and/or translational repression (Krol et al., 2010).

Altered miRNA expression is directly associated with cancer initiation, progression, and metastasis and is observed in a wide variety of human malignancies (Di Leva and Croce, 2010). The let-7 miRNA family members act as tumor suppressors by inhibiting expression of oncogenes and key regulators of mitogenic pathways including RAS, MYC, and HMGA2 (Büssing et al., 2008). let-7 is downregulated in numerous different cancers, and low let-7 correlates with poor prognosis (Boyerinas et al., 2010; Shell et al., 2007; Takamizawa et al., 2004). Restoration of let-7 expression effectively inhibits cancer growth in mouse models of lung and breast cancers (Barh et al., 2010; Esquela-Kerscher et al., 2008; Slack, 2009; Trang et al., 2010; Yu et al., 2007a). In humans, there are 12 let-7 family members (let-7a-1, -2, -3; let-7b; let-7c; let-7d; let-7e; let-7f-1, -2; let-7g; let-7i; miR-98) located at eight different chromosomal loci. Of note, many tumors are characterized by the coordinate downregulation of multiple let-7 miRNAs (Shell et al., 2007).

The developmentally regulated RNA-binding protein Lin28 was found to selectively repress expression of let-7 miRNAs (Heo et al., 2008; Newman et al., 2008; Rybak et al., 2008; Viswanathan et al., 2008). This posttranscriptional regulation of let-7 by Lin28 is required for normal development and contributes to the pluripotent state by preventing let-7-mediated differentiation of embryonic stem cells (ESCs). Lin28 overexpression or let-7 inhibition with antisense RNAs promotes reprogramming of human and mouse fibroblasts to induced pluripotent stem cells (iPSCs) (Martinez and Gregory, 2010; Melton et al., 2010; Yu et al., 2007b). Unlike in *C. elegans* where a single *Lin28* gene is responsible for repression of let-7 expression and control of

developmental timing, the mammalian genome encodes two Lin28 paralogs, Lin28 (hereafter Lin28A) and Lin28B (Guo et al., 2006; Lehrbach et al., 2009; Moss et al., 1997; Van Wynsberghe et al., 2011; Viswanathan and Daley, 2010). Lin28B also represses expression of multiple let-7 members, and genome-wide association studies (GWAS) have linked *Lin28B* with the determination of human height and control of the age of onset of puberty and menopause, phenotypes that are recapitulated in a mouse model (Zhu et al., 2010). Activation of Lin28A/Lin28B occurs in several different primary human tumors, and these tumors display low levels of let-7 expression (Iliopoulos et al., 2009; Viswanathan et al., 2009). Indeed, Lin28A/Lin28B function as oncogenes that promote cellular transformation when ectopically expressed (Iliopoulos et al., 2009; Viswanathan et al., 2009; West et al., 2009). Importantly, this effect is abrogated when let-7 is reintroduced into these cells (Iliopoulos et al., 2009; Viswanathan et al., 2009). Therefore, Lin28-mediated cellular transformation is directly dependent on let-7 levels. Conversely, depletion of Lin28A or Lin28B in human cancer cells results in decreased cell proliferation (Chang et al., 2009; Iliopoulos et al., 2009; Viswanathan et al., 2009). Lin28A/Lin28B may contribute to the development of aggressive, poorly differentiated tumors because their expression is associated with advanced disease in hepatocellular carcinoma (HCC), chronic myeloid leukemia (CML), Wilms' tumor, ovarian carcinoma, colon adenocarcinoma, and germ cell tumors (Dangi-Garimella et al., 2009; Guo et al., 2006; Iliopoulos et al., 2009; Ji and Wang, 2010; King et al., 2011; Liang et al., 2010; Lu et al., 2009; Oh et al., 2010; Peng et al., 2010; Viswanathan et al., 2009; Wang et al., 2010; West et al., 2009; Yang et al., 2010) and is associated with poor clinical outcome and patient survival in HCC, colon, and ovarian cancer (King et al., 2011; Lu et al., 2009; Viswanathan et al., 2009). In the case of Lin28B, rare amplification or translocation events might explain activation in some cases (Viswanathan et al., 2009). A more common mechanism might be transcriptional activation by upstream factors. For example, c-Myc binds to both *Lin28A* and *Lin28B* loci and activates expression of these genes (Chang et al., 2009). In a breast cancer model, transient expression of Src oncoprotein results in a transformed cell line that forms self-renewing mammospheres harboring tumor-initiating cells (Iliopoulos et al., 2009). The transformation process involves NF-κB activation leading to direct transcriptional upregulation of Lin28B, consequent let-7 loss, and derepression of the let-7 target gene IL-6. Because IL-6 activates NF-κB, this regulatory circuit represents a positive feedback loop, providing a molecular link between inflammation and cancer.

Selective regulation of let-7 expression involves Lin28A binding to the terminal loop of let-7 precursors, a molecular recognition that requires both the cold shock domain (CSD) and CCHC-type zinc finger RNA-binding domains of the Lin28A protein (Piskounova et al., 2008). Lin28A recruits the activity of a terminal uridylyltransferase (TUTase), Zcchc11 (also known as TUTase4 or TUT4), that inhibits pre-let-7 processing by Dicer and leads to the rapid decay of oligouridylated pre-let-7 RNAs (Hagan et al., 2009; Heo et al., 2009). Although both Lin28A and Lin28B can recruit Zcchc11/TUT4 to uridylate pre-let-7 in vitro, the molecular mechanism of the Lin28B-mediated blockade of let-7 expression has yet to be determined (Heo et al., 2008, 2009). Here we investigate the regulation of let-7 expression by Lin28B. Surprisingly, we find that despite their high degree of homology, Lin28A and Lin28B function through distinct mechanisms. Depletion of Zcchc11 affects let-7 expression only in Lin28A-expressing cancer cells, whereas Lin28B functions through a Zcchc11-independent mechanism. We find that Lin28A and Lin28B are differentially localized in cells with predominantly cytoplasmic Lin28A, whereas due to its functional nuclear localization signals, Lin28B accumulates in the nucleus, where it binds pri-let-7 miRNAs to block processing by the Microprocessor. In contrast, Lin28A functions in the cytoplasm by blocking at the Dicer step and recruiting the TUTase to uridylate pre-let-7. Our findings identify Zcchc11 as a possible therapeutic target in Lin28A-expressing cancers. Accordingly, we demonstrate that Zcchc11 depletion selectively inhibits the tumorigenic capacity and metastatic potential of Lin28A- but not Lin28B-expresing human cancer cells and xenografts. Our results illuminate the distinct mechanisms by which Lin28A and Lin28B function and have broad implications for the development of new strategies for cancer therapy.

RESULTS

Lin28B Regulates let-7 Expression through a Zcchc11-Independent Mechanism

The paralogous RNA-binding proteins Lin28A and Lin28B have a high degree of sequence identity and conserved domain organization (Figure 1A), and both proteins selectively block let-7 expression (Newman et al., 2008; Viswanathan et al., 2008). We screened several human cancer cell lines and found that some express Lin28A, whereas others express Lin28B (Figure 1B). We did not observe coexpression of both Lin28A and Lin28B in any cell line, suggesting that their expression may be mutually exclusive. We found ubiquitous Zcchc11 expression. HeLa cells express Zcchc11 but neither Lin28A nor Lin28B. Because Lin28A-mediated repression of let-7 in mouse ESCs (mESCs) involves the TUTase Zcchc11, we next asked whether Lin28A and Lin28B function through the same mechanism to block let-7 processing. Previous reports have used recombinant Lin28A and Lin28B interchangeably in biochemical assays, demonstrating that Lin28B is capable of enhancing Zcchc11 activity in vitro; however, the physiological relevance of these observations remains unknown (Heo et al., 2009).

To begin to investigate whether both Lin28A and Lin28B function through a Zcchc11 TUTase-dependent mechanism, we performed coimmunoprecipitation (co-IP) experiments. Myc-tagged Lin28A, Lin28B, or Ago2 were coexpressed with either Flag-tagged Zcchc11 or Flag-EIF6 control (Figure 1C). Because the Lin28A-Zcchc11 interaction has been shown to be RNA dependent, we also coexpressed pri-let-7g (Heo et al., 2009). Consistent with earlier reports, myc-Lin28A was found to be associated with affinity-purified Flag-Zcchc11 (Heo et al., 2009). However, we were unable to detect a physical interaction between myc-Lin28B and Flag-Zcchc11. We performed additional co-IP experiments in which we titrated the amount of exogenously expressed Flag-Zcchc11. These experiments confirmed the specific physical interaction of Zcchc11 and Lin28A,

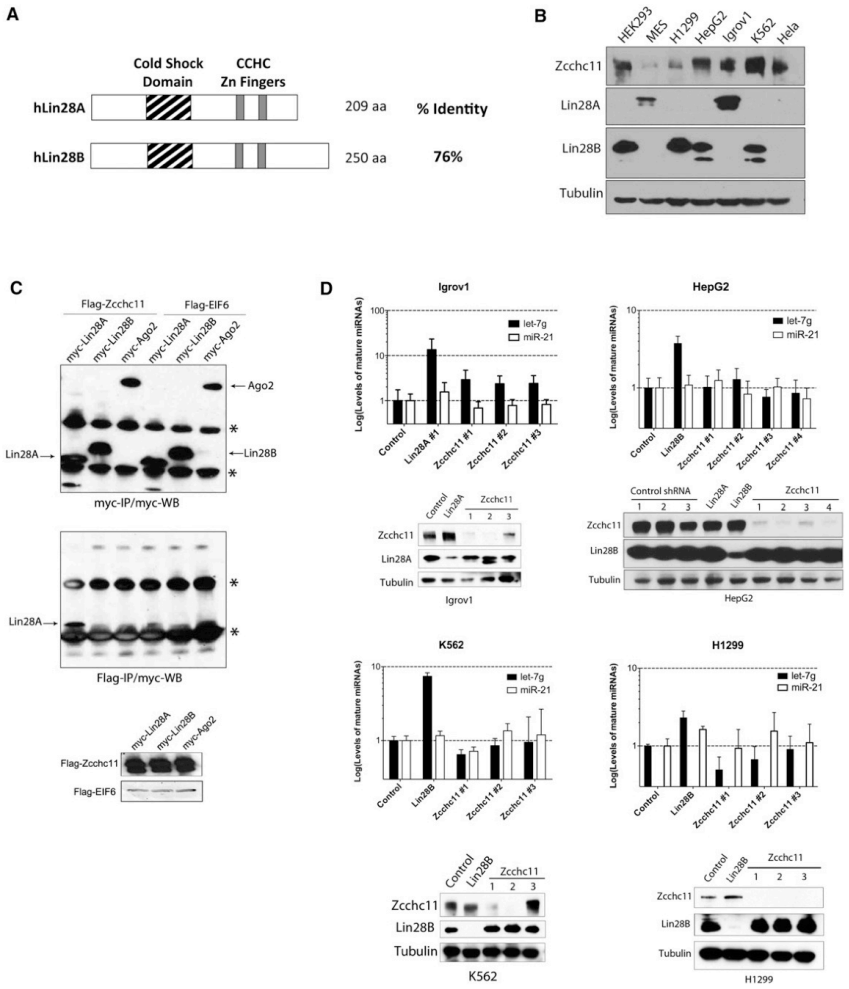


Figure 1. Lin28B Regulates let-7 Biogenesis through a Zcchc11-Independent Mechanism

(A) Schematic representation of human Lin28A and Lin28B.

(B) Western blot analysis of Zcchc11, Lin28A, and Lin28B in extracts prepared from human cancer cell lines.

(C) Coimmunoprecipitation (co-IP): HeLa cells were cotransfected with human myc-Lin28A, myc-Lin28B, or myc-Ago2 with either Flag-Zcchc11 or Flag-EIF6. Flag-IP and Flag- and Myc-western blots were performed to detect expression and interaction, respectively. See also Figure S1.

(D) Stable knockdown of Zcchc11 leads to upregulation of mature let-7g levels in Lin28A-expressing cells but not Lin28B-expressing cell lines. miRNA levels were measured by qRT-PCR. Error bars represent SEM (n = 3). Protein knockdown was monitored by western blot.

whereas myc-Lin28B was not detected in any of the Flag-Zcchc11 IPs (Figure S1A available online). This was additionally confirmed by the co-IP of endogenous Lin28A in Igrov1 cells (Figure S1B). Together, these results indicate that unlike for Lin28A, we could not detect any physical interaction between Lin28B and Zcchc11.

Next, to address the functional requirement of Zcchc11 in the Lin28A- and Lin28B-mediated repression of let-7 expression, we

performed a series of knockdown experiments to deplete Zcchc11 in a panel of human cancer cell lines. We used shRNAs to deplete Lin28A or Zcchc11 expression in Igrov1 cells and measured the effect on let-7 expression by quantitative reverse transcription PCR (q.RT-PCR). As expected, depletion of Lin28A led to an ~10-fold increase in let-7 levels. Knockdown of Zcchc11 with three independent shRNAs also led to elevated mature let-7 levels (Figure 1D). Therefore, Zcchc11 is involved in the repression of let-7 expression in this Lin28A-expressing human cancer cell line as has been previously reported in mESCs and embryonal carcinoma cells (Hagan et al., 2009; Heo et al., 2009). We performed analogous experiments in three different Lin28B-expressing cancer cell lines, HepG2, K562, and H1299 (Figure 1D), and found no significant effect on mature let-7 levels in any of the cell lines when Zcchc11 was depleted. In contrast, knockdown of Lin28B consistently led to the expected increase in mature let-7. Overall our results indicate that Zcchc11 negatively regulates let-7 expression in Lin28A- but not Lin28B-expressing cell lines, suggesting that Lin28B employs a Zcchc11-independent mechanism to block let-7 processing.

Lin28B Localizes to the Nucleus

We sought potential explanations for the functional differences between Lin28A and Lin28B. We used immunofluorescence assays to examine the subcellular localization of the endogenous Lin28A and Lin28B proteins (Figure 2A). Lin28A was mostly localized to the cytoplasm of Igrov1 cells, whereas Lin28B localized to specific foci in the nuclei of H1299 cells where it colocalized with the nucleolar marker Fibrillarin. To further confirm the localization of Lin28B in the nucleoli, we performed immunofluorescence assays on H1299 cells in which Lin28B expression (or control) was stably knocked down by shRNA and showed that the observed nucleolar staining pattern is specific to Lin28B (Figure 2B). These data were further confirmed by biochemical fractionation and western blot of both cell lines (Figure 2C). Consistent with published data, we found Zcchc11 only in the cytoplasmic fraction in both the Lin28A- and Lin28B-expressing cell lines (Figure 2C). These data suggest that the divergence in the mechanisms by which Lin28A and Lin28B block let-7 biogenesis derives from their differential subcellular localization. The lack of physical and functional interactions between Zcchc11 and Lin28B is therefore likely due to their localization to distinct cellular compartments, even though recombinant Lin28B has the capacity to enhance Zcchc11 activity in vitro (Heo et al., 2009).

Lin28B Contains Functional Nuclear Localization Signals

Lin28B protein has an extended C terminus compared to Lin28A, which upon closer inspection contains a putative bipartite nuclear localization signal (NLS), KK[GPSVQ]KRKK. Another potential NLS, RRPK[GKTLQ]KRKPK, was identified in the linker region that connects the two functional RNA-binding domains (Figure 2D). To test the function of these putative NLS, we generated constructs for the expression of a series of GFP fusion proteins. We transiently transfected HeLa cells with these constructs and analyzed the subcellular localization of the GFP fusion proteins by microscopy (Figure 2E). Consistent with the localization of endogenous Lin28A in Igrov1 cells, we found that Lin28A-GFP localized mainly to the cytoplasm. Lin28B-GFP predominantly localized to specific foci in the nucleus, again recapitulating the nucleolar localization of endogenous Lin28B observed in H1299 cells. When we exogenously expressed the Lin28BΔNLS#1 truncation, we observed increased signal in the cytoplasm; however, some nucleolar localization still remained consistent with the presence of a second NLS. Indeed, the double-mutant Lin28B-GFP protein lacking both NLS showed cellular localization similar to that of GFP alone, suggesting that both NLS elements are important for nuclear and nucleolar localization of Lin28B (Figure 2E). To further determine whether these sequences represent functional NLS, we examined the localization of the NLS#1-GFP and NLS#2-GFP (Figure 2E). When exogenously expressed in HeLa cells, NLS#1-GFP localized nearly entirely throughout the nucleus including the nucleoli. This is in contrast to the control GFP construct that is broadly distributed throughout the cell. NLS#2-GFP was nearly entirely localized to the nucleoli (Figure 2E). Together these results identify that NLS#1 amino acid sequence represents a bona fide NLS, and that NLS#2 is a functional nucleolar localization signal (NoLS). NoLS properties are less well known and have only recently been studied at the amino acid sequence level (Scott et al., 2010). Several reports suggest that proteins with a NoLS also contain an NLS that allows them to cross the nuclear membrane before localizing to the nucleoli. These GFP-fusion results were confirmed by biochemical fractionation of HeLa cells transiently expressing either Flag-Lin28A, full-length, truncated, or double-mutant Flag-Lin28B. As expected Flag-Lin28A was predominantly present in the cytoplasmic fraction, Flag-Lin28B was mostly nuclear, NLS#1 mutant only showed marginal increase of signal in the cytoplasmic fraction, whereas the double-mutant Flag-Lin28B showed a more significant increase in cytoplasmic signal (Figure 2F).

Lin28B Localizes to Nucleoli where the Microprocessor Is Absent

We next compared the specific localization of Lin28B with that of the nuclear miRNA processing machinery. We found that the Microprocessor components DGCR8 and Drosha colocalize in the nucleoplasm but are excluded from nucleoli (Figure 3A). To further confirm that localization of Lin28B is distinct and nonoverlapping with the Microprocessor, we examined the colocalization of mCherry-DGCR8 and GFP-Lin28A/B proteins in transfected cells (Figure 3B). Lin28A localized mostly to the cytoplasm and therefore showed no overlap with the nuclear DGCR8. Lin28B localized to nucleoli and did not overlap with DGCR8 either. In contrast, the localization of the Lin28B NLS/NoLS mutant showed a broadly dispersed localization throughout the nucleus and cytoplasm (similar to GFP control) and displayed colocalization with DGCR8 in the nucleoplasm. We confirmed this finding that Lin28B and the Microprocessor normally occupy distinct compartments in the nucleus by performing large-scale biochemical fractionation and western blot of a stable HeLa cell line expressing Flag-Lin28B. Lin28B was specifically present in the nucleolar-enriched fractions whereas DGCR8 was only detectable in the nuclear fraction and not in the nucleolar

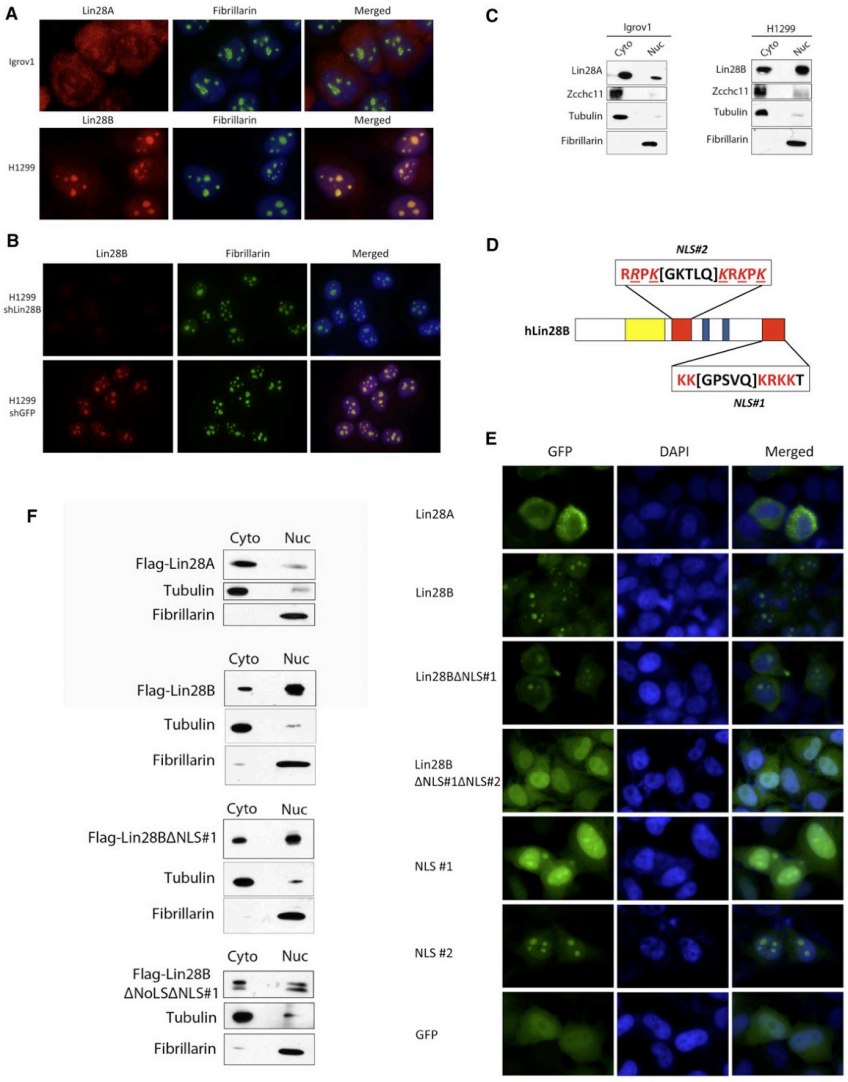


Figure 2. Lin28A and Lin28B Are Differentially Localized within the Cell

(A) Immunofluorescence detection of endogenous Lin28A in Igrov1 and Lin28B in H1299 cell lines. Fibrillarin, a known nucleolar protein, was used as a positive control.

(B) Immunofluorescence analysis of control and Lin28B knockdown H1299 cell lines.

(C) Biochemical fractionation of Igrov1 and H1299 cell lines. Endogenous Lin28A, Lin28B, and Zcchc11 in each fraction were detected by western blot. Fibrillarin was used as a nuclear marker; Tubulin was used as a cytoplasmic marker.

(D) Schematic of NLS in the Lin28B protein. An arginine as well as several lysines that were replaced by glycines are underlined and italicized.

(E) Localization of GFP-Lin28 fusion proteins in HeLa cells.

(F) Fractionation of Flag-Lin28 proteins, exogenously expressed in HeLa cells. Proteins were detected by Flag-western blot.

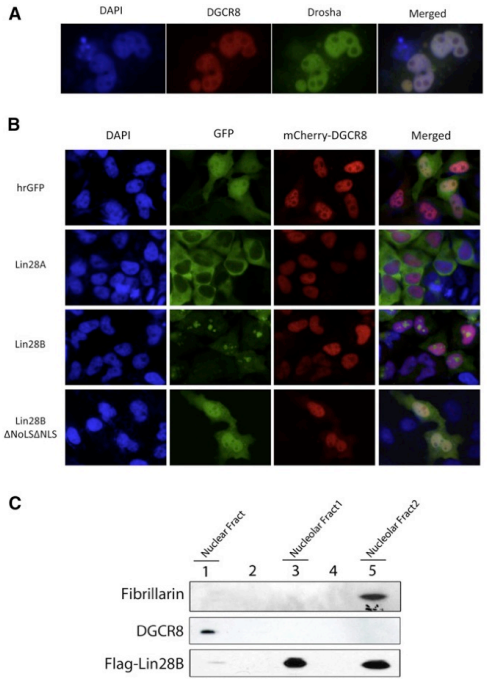


Figure 3. Lin28B Localizes to Nucleoli where the Microprocessor Is Absent

(A) Colocalization of the Microprocessor components GFP-Drosha and mCherry-DGCR8 in HeLa cells reveals their distribution throughout the nucleus and exclusion from nucleoli.

(B) Localization of GFP-Lin28A, Lin28B, or mutant Lin28B proteins with mCherry-DGCR8 in HeLa cells reveals nonoverlapping localization of Lin28B and DGCR8.

(C) Fractionation of a Flag-Lin28B HeLa stable cell line. Flag-Lin28B and endogenous DGCR8 were detected by western blot and show a nonoverlapping subcellular localization of Lin28B and the Microprocessor. Fibrillarin was used as a control for nucleolar localization.

fractions (Figure 3C). Overall, these findings suggest a possible mechanism by which Lin28B blocks let-7 processing in the nucleus by sequestering pri-let-7 miRNAs in the nucleoli away from the Microprocessor.

Lin28B Directly Binds and Sequesters Pri-let-7

To further dissect the mechanism of the Lin28B-mediated let-7 processing block, we first compared the relative abilities of recombinant human Lin28A and Lin28B proteins to bind pre-let-7 (Figures 4A, 4B, and S2). We performed electromobility shift assay (EMSA) with pre-let-7g to analyze the relative binding affinities of the two recombinant proteins. We found that Lin28A and Lin28B have apparent dissociation constants (K_D) of approximately 0.6 nM and 0.5 nM, respectively (Figure 4B). Both these estimated K_D are much lower than previously reported for recombinant mouse Lin28A. This difference is likely due to the omission here of nonspecific yeast transfer RNA (tRNA) competitor used previously in the binding buffer (Piskounova et al., 2008). We also performed EMSA with pri-let-7g and demonstrated that both recombinant Lin28A and Lin28B are able to bind pri-let-7g with similar affinities (Figure 4C). Collectively, these assays reveal that both Lin28 proteins can directly bind to let-7 precursors with high affinity in vitro.

To gain further support for our model in which Lin28B binds and sequesters pri-let-7 in the nucleus to inhibit the Microprocessor, we next examined the RNA associated with Lin28B. We individually expressed and purified Flag-Lin28A and Flag-Lin28B, extracted the associated RNA, and analyzed relative levels of pri-let-7g by qRT-PCR. This RNA immunoprecipitation (RIP) analysis revealed that Lin28B directly associates with pri-let-7g RNA (Figure 4D). We detected an ~18-fold enrichment of pri-let-7 associated with Lin28B. Furthermore, we found substantially more pri-let-7 associated with Lin28B than with Lin28A, which is consistent with the differential subcellular localization of these proteins. Taken together, these results indicate that this preferential association of pri-let-7g with Lin28B likely reflects the distinct mechanism by which Lin28B represses let-7 expression rather than any possible intrinsic differences in the relative RNA-binding affinities of Lin28A and Lin28B proteins. Next, we examined by qRT-PCR the effect of transient Lin28B overexpression on pri-let-7g levels. Transient Lin28B overexpression led to ~12-fold accumulation of pri-let-7g levels (Figure 4E). In contrast, overexpression of Lin28A only had a more modest effect on pri-let-7g levels, consistent with its predominantly cytoplasmic localization. In order to further assess the effects of Lin28B overexpression on both pri- and mature miRNA levels, we utilized a Flag-Lin28B expressing a HeLa stable cell line. Analysis of several pri-miRNAs by qRT-PCR in this cell line demonstrated a substantial accumulation of pri-let-7 miRNAs, >10-fold for pri-let-7g and >3-fold for pri-let-7a-1. There was, however, no effect on levels of pri-miR-21 (Figure 4F). We detected a corresponding decrease in the levels of mature let-7, with >90% decrease for mature let-7g and ~40% decrease for mature let-7a. Again no effect was observed on levels of mature miR-21 (Figure 4F). Together, these data support our model whereby nuclear Lin28B directly associates with pri-let-7, sequestering it from cleavage by the Microprocessor to selectively inhibit let-7 maturation, and underscore our findings that the paralogous RNA-binding proteins Lin28A and Lin28B operate by distinct mechanisms to selectively repress let-7 expression.

Zcchc11 Inhibition Blocks the Tumorigenicity and Invasiveness of Lin28A- but Not Lin28B-Expressing Breast Cancer Cells In Vitro and In Vivo

We were motivated to further explore the relevance of our findings that Lin28A and Lin28B block let-7 processing through distinct mechanisms and to examine the effect of Zcchc11 inhibition on the tumorigenicity and invasiveness of human Lin28A/B-expressing cancer cells. Initially, we tested the consequences of Zcchc11 inhibition in the MCF10A ER-Src-inducible model of cellular transformation, where MCF10A-immortalized

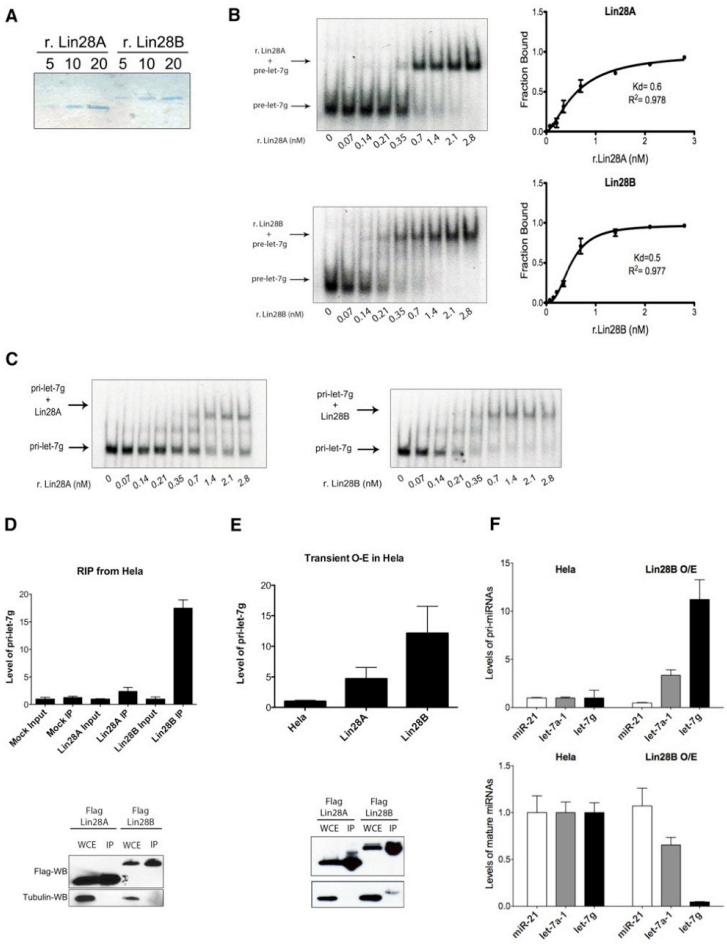


Figure 4. Lin28B Directly Binds and Sequesters Pri-let-7

(A) Colloidal blue staining of purified recombinant His-Lin28A and His-Lin28B proteins.

(B) Binding of r.Lin28A and r.Lin28B to pre-let-7g was assessed by EMSA performed with 0.5 nM 5′-end labeled pre-let-7g RNA and the indicated concentration of recombinant protein. Band intensities were quantitated from three independent experiments and represented as the fraction of bound pre-let-7g RNA in the plots. Values are given as average ± SEM (n = 3). See also Figure S2.

(C) EMSA performed indicated concentration of r.Lin28A and r.Lin28B with in vitro transcribed uniformly labeled pri-let-7g.

(D) RIP analysis of RNA associated with immunopurified Flag-Lin28A and Flag-Lin28B from HeLa cells. RNA was extracted from IP material and analyzed by qRT-PCR. Error bars ± SEM (n = 3). Lower panel indicates relative Lin28A and Lin28B expression levels by Flag-western blot.

(E) Accumulation of pri-let-7 by transient Lin28B expression in HeLa cell detected by qRT-PCR. Error bars ± SEM (n = 3). Lower panel indicates relative expression levels of Lin28A and Lin28B proteins detected by Flag-western blot in transfected cells.

(F) pri-let-7 accumulates (top panel) and mature let-7 levels decrease (bottom panel) in HeLa cells stably overexpressing Lin28B. Error bars ± SEM (n = 3).

breast epithelial cells become transformed 36 hr post-tamoxifen (TAM) treatment (Iliopoulos et al., 2009). We previously reported that Lin28B expression is activated and required for induction and maintenance of the transformed phenotype in this model (Figures S3A and S3B). In contrast, Zcchc11 depletion did not affect transformation or colony formation (Figures S3B–S3D)

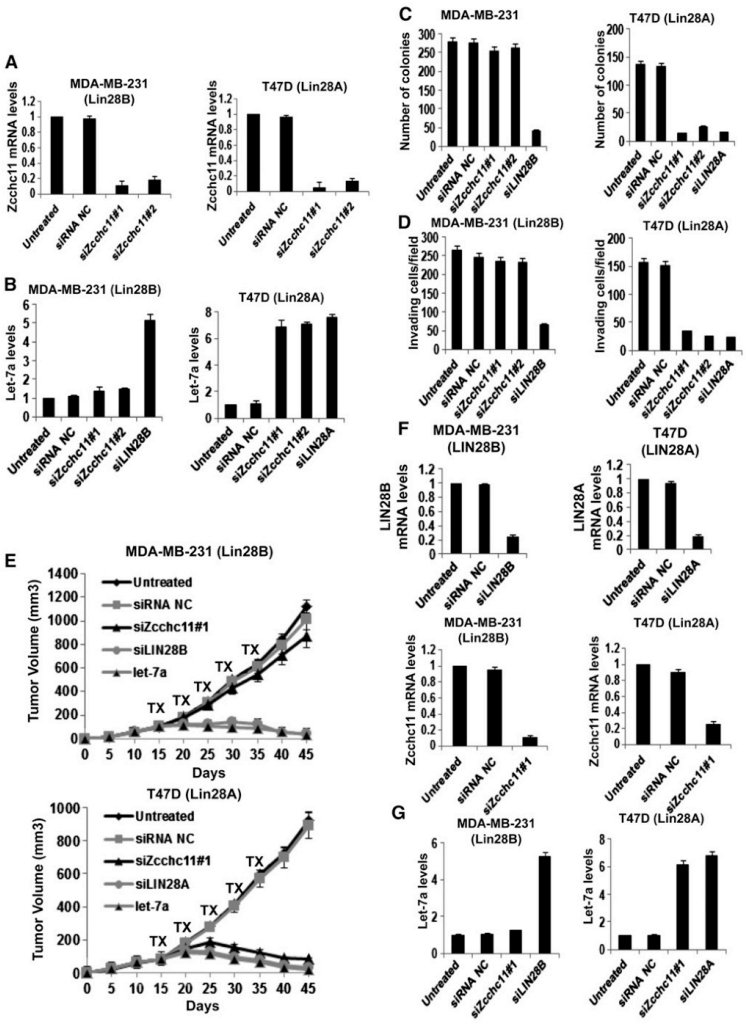


Figure 5. Zcchc11 Inhibition Blocks Tumorigenicity and Invasiveness of Lin28A-Expressing Breast Cancer Cells

(A) qRT-PCR analysis of Zcchc11 knockdown in MDA-MB-231 and T47D breast cancer cells. Error bars ± SEM (n = 3).

(B) Inhibition of Zcchc11 expression does not affect let-7a expression in Lin28B-expressing cells (MDA-MB-231), whereas it upregulates let-7a expression in Lin28A-expressing cells (T47D). Let-7a expression levels were measured by qRT-PCR in cells treated with siRNAs for 48 hr. Error bars ± SEM (n = 3).

(C) Inhibition of Zcchc11 expression did not affect the colony formation ability of MDA-MB-231 cells but suppressed the colony formation ability of T47D cells. The number of colonies was evaluated 20 days post-plating in soft agar. The experiment was repeated thrice, and the statistical significance was calculated using Student's t test.

(D) Inhibition of Zcchc11 expression did not affect the invasiveness of MDA-MB-231 cells but suppressed the invasive ability of T47D cells. The number of invasive cells was measured 16 hr post-transfection with indicated siRNAs. In all assays, 10 fields per insert were scored, and SD was measured. The experiment was repeated thrice, and the statistical significance was calculated using Student's t test.

(E) Inhibition of Zcchc11 expression did not suppress tumor growth of MDA-MB-231 cells in xenografts, but it inhibited tumor growth of T47D cells in xenografts. The treatments with indicated siRNA were performed i.p. for five cycles starting on day 15. Each treatment group consisted of five mice.

and did not affect the expression of let-7a and its direct downstream target IL-6 (Figures S3D–S3F). Accordingly, Zcchc11 inhibition did not affect the tumorigenicity of MCF10A ER-Src-transformed cells in xenograft assays (Figure S3G). Overall, these data suggest that inhibition of Zcchc11 does not have an inhibitory effect on the tumorigenicity and invasiveness of Lin28B-expressing MCF10A ER-Src cells.

To further explore the distinct requirements for Zcchc11 in Lin28A- and Lin28B-expressing cancer, we compared the effects of Zcchc11 inhibition on the tumorigenicity and invasiveness of MDA-MB-231 breast cancer cells (Lin28B-expressing cells) relative to T47D breast cancer cells (Lin28A-expressing cells). We found that suppression of Zcchc11 expression did not affect let-7a expression in MDA-MB-231 cells but led to a 7-fold increase in mature let-7a levels in T47D cells (Figures 5A and 5B). Furthermore, Zcchc11 inhibition did not affect the tumorigenicity and invasiveness of MDA-MB-231 cells, although it suppressed both the colony formation ability and invasiveness of T47D cells (Figures 5C and 5D). Zcchc11 inhibition had similar effects on the tumor growth of these cell lines in xenografts. Specifically, Zcchc11 knockdown did not affect tumor growth of MDA-MB-231, whereas it suppressed T47D tumor growth (Figure 5E). Synthetic let-7a miRNA suppressed both MDA-MB-231 and T47D tumor growth (Figure 5E). Also, in the tumors derived from MDA-MB-231 xenografts (day 30), let-7a expression was not affected by inhibition of Zcchc11, whereas Lin28B suppression increased let-7a levels about 5-fold (Figures 5F and 5G). On the other hand, both Zcchc11 and Lin28A inhibition resulted in upregulation of let-7a expression to similar levels in T47D-derived tumors (day 30) (Figures 5F and 5G).

In addition to the breast cancer cells, we tested the effects of Zcchc11 inhibition on tumor growth of several other (liver, lung, ovarian, melanoma, colon) cancer cell types (Figure 6A). As above, we found that Zcchc11 inhibition (Figure 6B) blocked the growth of Lin28A-expressing tumors (Igrov1) and did not affect the growth of Lin28B-expressing tumors (HepG2, H1299, SK_MEL_28, CaCO2) (Figure 6A). Lin28A and Lin28B inhibition suppressed the growth of the corresponding tumors (Figures 6C and 6D). Taken together, these data suggest that Zcchc11 plays a role in the tumorigenicity and invasiveness of Lin28A-expressing cancer cells, but depletion of Zcchc11 in Lin28B-expressing cancer cell lines has no effect on cancer growth.

Lin28A and Lin28B Expression Levels in Human Colon and Breast Tissues

To further study the disease relevance of our findings, we measured Lin28A and Lin28B expression in human colon and breast tissues. We found that Lin28A or Lin28B is upregulated whereas let-7a is downregulated in colon adenocarcinomas relative to normal colon tissues (Figure 7A). Specifically, we identified that Lin28A was upregulated in 19/45 colon adenocarcinomas, whereas Lin28B was upregulated in 14/45 colon adenocarcinomas. We found that the colon tumors with Lin28A overexpression had very low levels of Lin28B and vice versa. Furthermore, immunohistochemistry and in situ hybridization analyses in normal and colon cancer tissues revealed that Lin28A or Lin28B proteins are upregulated whereas let-7a is downregulated in colon carcinomas relative to normal colon tissues (Figures 7B and S4). Similar to the mRNA data, immunohistochemistry revealed that tumors that expressed high Lin28A protein levels had low levels for Lin28B and vice versa. This is consistent with our analysis of human cancer cell lines where we did not find cells that express both Lin28A and Lin28B (Figure 1B and data not shown).

We also found that Lin28A or Lin28B was upregulated whereas let-7a was downregulated in breast cancer relative to normal breast tissues (Figure 7C). Specifically, we identified that Lin28A was upregulated in 9/33 breast carcinomas, whereas Lin28B was upregulated in 10/33 breast carcinomas. Similar to the colon tissues, we detected that the breast tumors that overexpressed Lin28A had very low levels of Lin28B and vice versa. In addition, we detected that Lin28A was significantly upregulated in HER2-overexpressing breast tumors, whereas Lin28B was significantly upregulated in triple-negative (ER^{-},PR^{-},$HER2^{-}$) breast tumors (Figure 7D). Furthermore, according to our previous studies, Lin28B expression is a part of an inflammatory circuit and is regulated by NF-κB transcription factor (Iliopoulos et al., 2009). Thus, we tested how NF-κB activity correlates with Lin28A or Lin28B mRNA levels in human breast tumors. We found a statistically significant correlation between NF-κB nuclear levels and Lin28B but not Lin28A expression levels (Figure 7E). These data suggest that NF-κB regulates the Lin28B but not Lin28A pathway. Also, in order to have a broader view of Lin28A and Lin28B expression levels in human cancer tissues, we tested their expression levels in kidney, liver, lung, ovarian, prostate, and thyroid cancer (Figure 7F). These data reveal that Lin28B is upregulated in liver, ovarian, and thyroid carcinomas.

DISCUSSION

Lin28A and Lin28B inhibit let-7 miRNA biogenesis in ESCs and cancer, and it has been assumed that both proteins block let-7 expression through the same mechanism, by recruiting the TUTase Zcchc11 (TUT4) in the cell cytoplasm to uridylate pre-let-7 and target it for degradation. Moreover, these paralogous proteins have been used interchangeably in several in vitro assays (Heo et al., 2009). Our results provide evidence that despite their high degree of homology, Lin28A and Lin28B function through distinct mechanisms to block let-7 processing, a finding that has broad implications for the development

(F) qRT-PCR analysis of siRNA inhibition of Lin28B, Lin28A, and Zcchc11 in xenograft tumors (day 30) derived from MDA-MB-231 and T47D cells. Error bars ± SEM (n = 3).

(G) Inhibition of Lin28B but not of Zcchc11 allows upregulation of let-7a expression levels in MDA-MB-231 xenograft tumors (day 30). However, inhibition of Lin28A or Zcchc11 results in let-7a upregulation in T47D xenograft tumors. Let-7a expression levels were measured by qRT-PCR on tumors untreated or treated with indicated siRNA. Error bars ± SEM (n = 3). See also Figure S3.

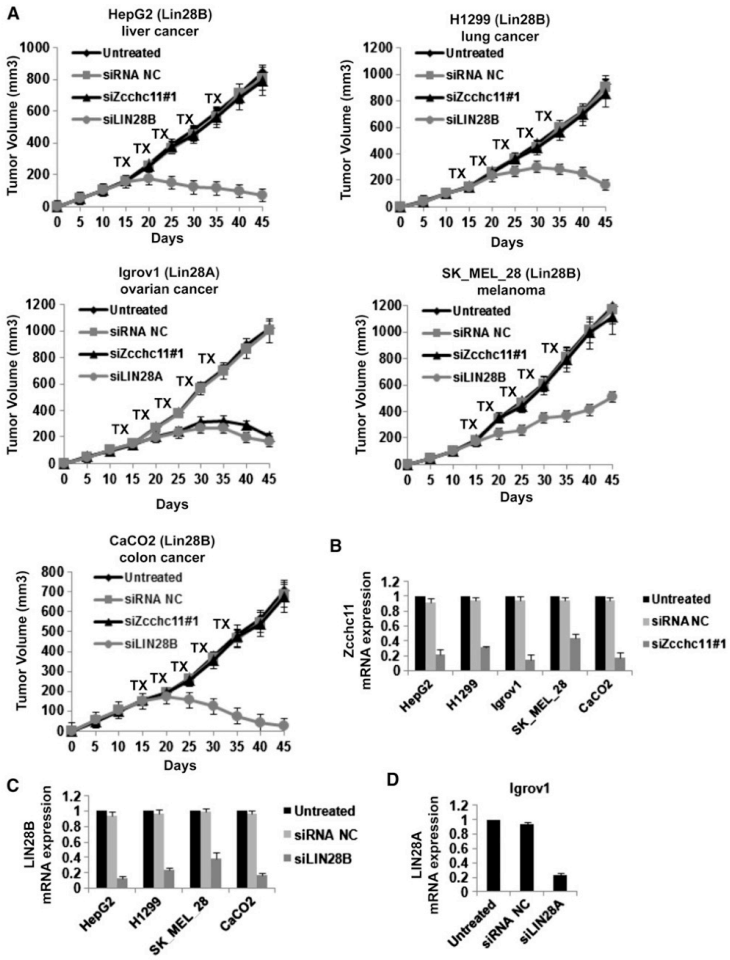


Figure 6. Inhibition of Zcchc11 Expression Suppresses Tumor Growth of Lin28A- but Not Lin28B-Expressing Xenografts

(A) Xenograft experiments were performed with a variety of different human cancer cell lines. Mice were treated with the indicated siRNA for five cycles starting on day 15. For all cells lines tested, each treatment group consisted of five mice. Although inhibition of Lin28A or Lin28B suppressed tumor growth in the relevant xenografts, inhibition of Zcchc11 inhibited growth only of Lin28A- but not Lin28B-expressing tumors. Error bars ± SEM (n = 3).

(B) Analysis of siRNA inhibition of Zcchc11 in xenograft tumors (day 30) derived from the indicated cells.

(C) Analysis of siRNA inhibition of Lin28B in xenograft tumors (day 30) derived from the indicated cells.

(D) Analysis of siRNA inhibition of Lin28A in xenograft tumors (day 30) derived from IGROV1 cells. mRNA expression levels were measured by qRT-PCR on tumors untreated or treated with the indicated siRNA. Error bars ± SEM (n = 3).

of new cancer therapies and the potential use of Zcchc11 inhibitors in Lin28A-expressing tumors but not in Lin28B-expressing tumors. This distinction derives from the differential subcellular localization of these two proteins: Lin28A localizes primarily to the cytoplasm, whereas Lin28B contains functional nuclear localization signals and specifically localizes to nucleoli.

Due to the differential subcellular localization of the two proteins, we find that in human cancer cell lines, Lin28A and Lin28B block let-7 processing at different steps of the

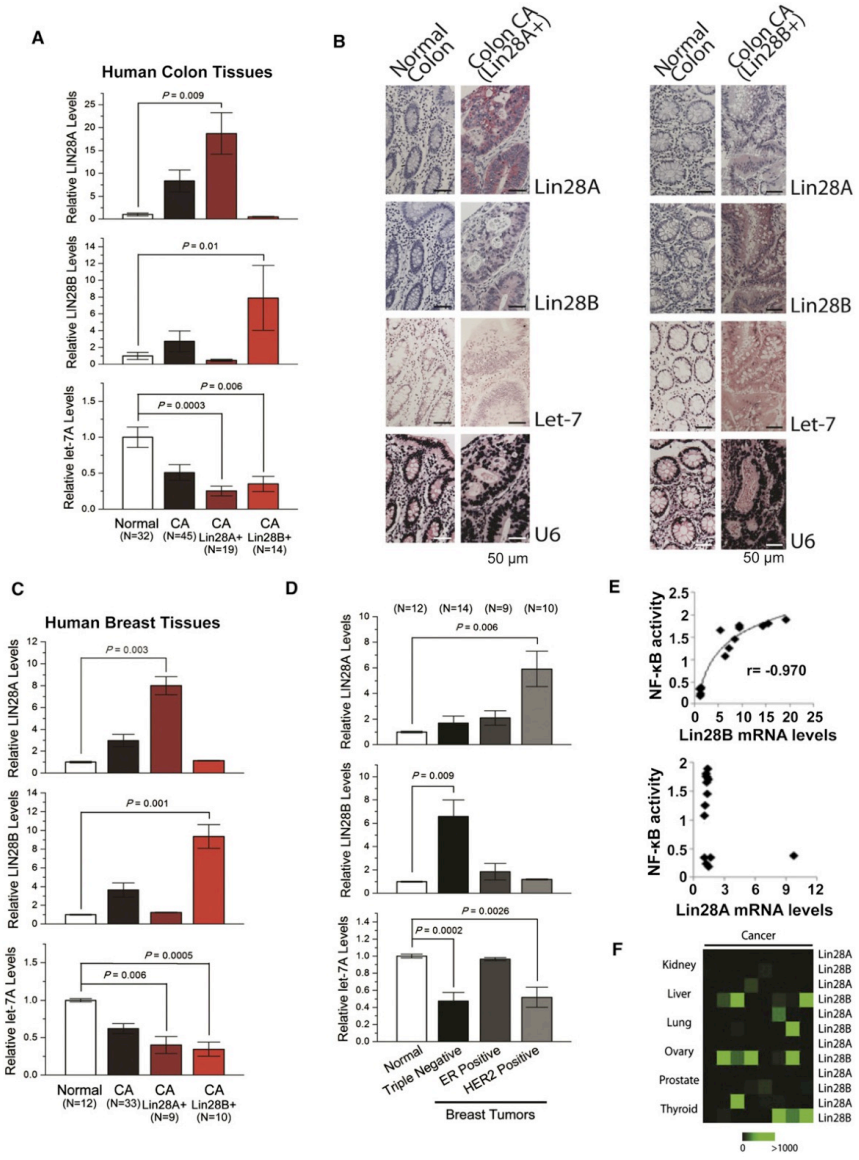


Figure 7. Lin28A and Lin28B Expression in Primary Human Cancers

(A) qRT-PCR analysis of Lin28A, Lin28B, and let-7a expression levels in normal and colon cancer tissues. Tumor samples were further classified into two groups expressing either high Lin28A or Lin28B. Data expressed as mean ± SEM (n = 3).

(B) Immunohistochemistry for Lin28A and Lin28B and in situ hybridization for let-7a and U6 in normal colon tissues and colon adenocarcinomas. See also Figure S4.

miRNA-processing pathway. However, the steps at which let-7 processing is blocked by Lin28 in various different organisms are controversial. A recently published report by Van Wynsberghe et al. proposes that Lin28 binds pri-let-7 and blocks let-7 expression cotranscriptionally in *C. elegans* and disputes earlier conclusions that Lin28 functions at steps further downstream in the let-7 biogenesis pathway (Lehrbach et al., 2009; Van Wynsberghe et al., 2011). They also report that a very small fraction of Lin28A in human ESCs (hESCs) localizes to the nucleus and binds pri-let-7 miRNAs, although pre-let-7 is bound abundantly (Van Wynsberghe et al., 2011). This is consistent with our results that demonstrate that a small fraction of Lin28A localizes to the nucleus in an Igrov1 cell line. We also find that Lin28A binds pri-let-7 miRNA, but not as much as Lin28B, and earlier reports have demonstrated that purified Lin28A can inhibit the Microprocessor in vitro (Newman et al., 2008; Viswanathan et al., 2008). Also, although our data demonstrate that Lin28B-mediated repression of let-7 expression is Zcchc11 (TUT4) independent in multiple different cell types, it remains possible that in certain contexts or cell types, Lin28B may localize to the cytoplasm and utilize Zcchc11/TUT4 to repress let-7 biogenesis. For example, uridylated pre-let-7 was previously detected in Huh7 cells, and Lin28B is reportedly localized to the cytoplasm in Huh7 cells (Guo et al., 2006; Heo et al., 2008).

Our demonstration that the Microprocessor is excluded from nucleoli raises the possibility that sequestration of certain pri-miRNAs to nucleoli by specific RNA-binding proteins may prove to be a more general strategy for the posttranscriptional control of miRNA biogenesis. Previous reports have demonstrated that nucleoli contain machinery responsible for modifying small nucleolar RNAs, for example through RNA methylation or 3′ uridylation. Whether these nucleolar mechanisms play a role in pri-miRNA regulation remains to be determined (Boisvert et al., 2007; Matera et al., 2007). It is possible that additional new factors may be involved in sequestering and possibly degrading and/or modifying pri-let-7 miRNAs in the nucleoli in a Lin28B-dependent manner. The identification of such factors could reveal new potential therapeutic targets aimed at restoring let-7 expression in Lin28B-expressing cancers.

Our proof-of-concept studies with human breast and ovarian cancer cell lines demonstrate that inhibition of Zccch11 may represent a possible new therapeutic target for cancer. Indeed knockdown of Zcchc11 effectively inhibits the tumorigenic capacity and metastatic potential of human breast and ovarian cancer cells and xenografts. Importantly, however, our data also predict that the therapeutic potential of Zcchc11 inhibition will be limited to Lin28A-expressing cancers. Although Lin28A expression is relatively uncommon in several human cancer cell lines, our analysis of primary human colon and breast cancer indicates that upregulation of Lin28A or Lin28B occurs in a large proportion of tumors with approximately equal frequency for each protein. Furthermore, expression of Lin28A or Lin28B seems to distinguish different classes of breast cancers (Figure 7). Therefore, the identification of small-molecule inhibitors of Zcchc11 may lead to the development of novel chemotherapies for Lin28A-expressing cancers.

EXPERIMENTAL PROCEDURES

Cloning

Myc-Lin28A and -Lin28B were cloned into pBK-EF1. Lin28A, Lin28B, and Lin28BΔNLS#1 were cloned into pFLAG-CMV2 vector (Sigma). Lin28BΔNoLSΔNLS#1 was generated by site-directed mutagenesis using the Quick Change kit (Stratagene). Lin28A, Lin28B, Lin28BΔNLS#1, and Lin28BΔNoLSΔNLS#1 were cloned into CT-GFP-Topo (Invitrogen). NLS#1 and NLS#2 oligos were annealed before ligating into CT-GFP-Topo. N-terminal Cherry-DGCR8 fusion construct was generated by subcloning Cherry cDNA into p3xFLAG-CMV14-DGCR8 (Gregory et al., 2004). Lin28A and Lin28B were subcloned into Pet21 for His-tagged recombinant protein expression. Pri-let-7g was previously reported (Viswanathan et al., 2008), Cloning primers are listed in Table S1.

Transfection, Knockdowns, and Stable Cell Line Generation

All transfections were preformed with Lipofectamine (Invitrogen) per manufacturer's instructions. shRNA constructs were generated using pLKO.1 Puro (Addgene #8453) and pLKO.1 Hygro (Addgene #24150) vectors. The sequences of the shRNA hairpins and siRNAs are listed in Table S2. Lentivirus production, infection, and stable cell line selection are as described (Viswanathan et al., 2009). siRNAs used in this study were as follows: siRNA negative control (100 nM), (cat no. AM4611, Ambion Inc); siRNA Zcchc11#1 (100 nM) (cat no. s23551, Ambion Inc); siRNA Zcchc11#2 (100 nM) (cat no. s23553, Ambion Inc); siRNA Lin28B (100 nM) (cat no. s52477, Ambion Inc); siRNA Lin28A (100 nM) (cat no. s36195, Ambion Inc).

Western Blotting

The following antibodies were used: Lin28A (Cell Signaling, 3978), Lin28B (Cell Signaling, 4196), Zcchc11 (Protein Tech Group, 18980-1-AP), Zcchc11 (Imgenex, IMX-3587), DGCR8 (Protein Tech Group, 10996-1-AP), Fibrillarin (Abcam, ab18380), and β-Tubulin (Abcam AB6046).

Subcellular Fractionation

Cellular fractionation was done using NE-PER Nuclear and Cytoplasmic Extraction Kits (Pierce) per manufacturer's instructions. Large-scale fractionation was performed as described in the Extended Experimental Procedures.

EMSA

EMSA with purified His-Lin28A/B was performed as described but without competitor yeast tRNA (Piskounova et al., 2008). Complexes were resolved on native 3.5% or 5% polyacrylamide gels and visualized by autoradiography. Band intensities of scanned gels were quantified using ImageJ software and used to calculate percentage of probe bound. Graph-Pad Prism was used to plot data. Percent active protein was determined using stoichiometric binding reactions as described in Ryder et al. (2008). Hills equation for specific binding with one site, $Y = Bmax^{*}X^{\wedge}h/(K_D^{\wedge}h + X^{\wedge}h)$, was used to calculate K_D.

Colony Formation Assay

MDA-MB-231 cells and T47D cells were transfected with different siRNAs for 48 hr. Triplicate samples of 10^5 cells from each cell line were mixed 4:1 (v/v) with 2.0% agarose in growth medium for a final concentration of 0.4% agarose. The cell mixture was plated on top of a solidified layer of 0.5% agarose in

(C) qRT-PCR analysis of Lin28A, Lin28B, and let-7a in human normal and breast cancer tissues. Tumor samples were further classified into two groups expressing either high Lin28A or Lin28B. Data expressed as mean ± SEM (n = 3).
(D) Lin28A, Lin28B, and let-7a expression levels in different breast cancer subtypes.
(E) Correlation between Lin28A and Lin28B mRNA levels assessed by qRT-PCR with NF-κB phosphorylation status assessed by ELISA assay.
(F) Heatmap representation of Lin28A and Lin28B in carcinomas of different origin measured by qRT-PCR.

growth medium. Cells were fed every 6 to 7 days with growth medium containing 0.4% agarose. The number of colonies was counted after 20 days.

Invasion Assays

We performed invasion assays in MDA-MB-231, and T47D breast cancer cells were transfected with different siRNAs for 16 hr. Invasion of matrigel was conducted by using standardized conditions with BDBioCoat growth factor-reduced MATRIGEL invasion chambers (PharMingen). Assays were conducted per manufacturer's protocol, using 10% FBS as chemoattractant. Noninvading cells on the top side of the membrane were removed, whereas invading cells were fixed and stained with 4′-6-diamidino-2-phenylindole (DAPI, Vector Laboratories Inc.), 16 hr post-seeding.

Mouse Experiments

Xenograft experiments with human cancer cell lines are described in detail the Extended Experimental Procedures. Briefly, cells were injected subcutaneously in the right flank of athymic nude mice. Tumor growth was monitored every 5 days. In Vivo Ready siRNAs (Ambion Inc.) were mixed with Invivofectamine 2.0 liposomes (Ambion Inc) and injected intraperitoneally (i.p.) in a volume of 100 μl at a dose of 5 mg/kg.

Human Tissues and RNAs

Thirty normal colon tissues and 45 colon adenocarcinomas were collected from the translational pathology core laboratory of the Department of Pathology at UCLA. All subjects gave informed consent, and the study was approved by the UCLA Institutional Review Board. RNAs from 12 normal mammary tissues and 33 breast cancer tissues were purchased from Origene Inc. The ER, PR, and HER2 status for each of these breast carcinomas was known. Additional RNAs were purchased from Origene (8 renal cell carcinomas, 8 hepatocellular carcinomas, 8 squamous cell lung carcinomas, 8 ovarial adenocarcinomas, 8 prostate adenocarcinomas, 8 papillary thyroid carcinomas).

In Situ Hybridization and Immunohistochemistry

Sections of the colon adenocarcinomas and adjacent uninvolved tissues were analyzed by in situ hybridization as described (Iliopoulos et al., 2009) with modification (Extended Experimental Procedures). Double-DIG labeled Mircury LNA Detection probe for the detection of hsa-let-7 (1800-15, Exiqon) was used. Protocol for immunohistochemistry is described in detail the Extended Experimental Procedures, with the following antibodies: Lin28A (3978, Cell Signaling Technology) and Lin28B (LS-B3423, LSBio).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Extended Experimental Procedures, four figures, and two tables and can be found with this article online at doi:10.1016/j.cell.2011.10.039.

ACKNOWLEDGMENTS

We thank Dr. Akihiko Yoshimura, Keio University, Japan, for the human Flag-Zcchc11 plasmid; Dr. Na Liu, Memorial Sloan Kettering Cancer Center, for the Myc-Ago2 plasmid; Dr. Bharat Ramratnam, Brown University, for the GFP-Drosha plasmid; and Dr. Ramin Shiekhattar, Wistar Institute, for the Flag-Eif6 expression plasmid. We thank Chiara Giacomelli, Dr. Saurabh Patel, and Dr. Maria Hatziapostolou for technical assistance. R.I.G. was supported by grants from the US National Institute of General Medical Sciences (NIGMS) (1R01GM086386-01A1), The Harvard Stem Cell Institute, and The American Cancer Society (121635-RSG-11-175-01-RMC). R.I.G. is a Pew Research Scholar. D.I. was supported by start-up funds from the Department of Cancer Immunology & AIDS at Dana-Farber Cancer Institute.

Received: March 8, 2011
Revised: July 25, 2011
Accepted: October 24, 2011
Published: November 23, 2011

REFERENCES

Barh, D., Malhotra, R., Ravi, B., and Sindhurani, P. (2010). MicroRNA let-7: an emerging next-generation cancer therapeutic. Curr. Oncol. *17*, 70–80.

Boisvert, F.M., van Koningsbruggen, S., Navascués, J., and Lamond, A.I. (2007). The multifunctional nucleolus. Nat. Rev. Mol. Cell Biol. *8*, 574–585.

Boyerinas, B., Park, S.M., Hau, A., Murmann, A.E., and Peter, M.E. (2010). The role of let-7 in cell differentiation and cancer. Endocr. Relat. Cancer *17*, F19–F36.

Büssing, I., Slack, F.J., and Grosshans, H. (2008). let-7 microRNAs in development, stem cells and cancer. Trends Mol. Med. *14*, 400–409.

Chang, T.C., Zeitels, L.R., Hwang, H.W., Chivukula, R.R., Wentzel, E.A., Dews, M., Jung, J., Gao, P., Dang, C.V., Beer, M.A., et al. (2009). Lin-28B transactivation is necessary for Myc-mediated let-7 repression and proliferation. Proc. Natl. Acad. Sci. USA *106*, 3384–3389.

Dangi-Garimella, S., Yun, J., Eves, E.M., Newman, M., Erkeland, S.J., Hammond, S.M., Minn, A.J., and Rosner, M.R. (2009). Raf kinase inhibitory protein suppresses a metastasis signalling cascade involving LIN28 and let-7. EMBO J. *28*, 347–358.

Denli, A.M., Tops, B.B., Plasterk, R.H., Ketting, R.F., and Hannon, G.J. (2004). Processing of primary microRNAs by the Microprocessor complex. Nature *432*, 231–235.

Di Leva, G., and Croce, C.M. (2010). Roles of small RNAs in tumor formation. Trends Mol. Med. *16*, 257–267.

Esquela-Kerscher, A., Trang, P., Wiggins, J.F., Patrawala, L., Cheng, A., Ford, L., Weidhaas, J.B., Brown, D., Bader, A.G., and Slack, F.J. (2008). The let-7 microRNA reduces tumor growth in mouse models of lung cancer. Cell Cycle *7*, 759–764.

Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235–240.

Gregory, R.I., Chendrimada, T.P., Cooch, N., and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. Cell *123*, 631–640.

Guo, Y., Chen, Y., Ito, H., Watanabe, A., Ge, X., Kodama, T., and Aburatani, H. (2006). Identification and characterization of lin-28 homolog B (LIN28B) in human hepatocellular carcinoma. Gene *384*, 51–61.

Hagan, J.P., Piskounova, E., and Gregory, R.I. (2009). Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. Nat. Struct. Mol. Biol. *16*, 1021–1025.

Heo, I., Joo, C., Cho, J., Ha, M., Han, J., and Kim, V.N. (2008). Lin28 mediates the terminal uridylation of let-7 precursor MicroRNA. Mol. Cell *32*, 276–284.

Heo, I., Joo, C., Kim, Y.K., Ha, M., Yoon, M.J., Cho, J., Yeom, K.H., Han, J., and Kim, V.N. (2009). TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. Cell *138*, 696–708.

Hutvágner, G., McLachlan, J., Pasquinelli, A.E., Bálint, E., Tuschl, T., and Zamore, P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. Science *293*, 834–838.

Iliopoulos, D., Hirsch, H.A., and Struhl, K. (2009). An epigenetic switch involving NF-kappaB, Lin28, Let-7 MicroRNA, and IL6 links inflammation to cell transformation. Cell *139*, 693–706.

Ji, J., and Wang, X.W. (2010). A Yin-Yang balancing act of the lin28/let-7 link in tumorigenesis. J. Hepatol. *53*, 974–975.

King, C.E., Cuatrecasas, M., Castells, A., Sepulveda, A.R., Lee, J.S., and Rustgi, A.K. (2011). LIN28B promotes colon cancer progression and metastasis. Cancer Res. *71*, 4260–4268.

Krol, J., Loedige, I., and Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. Nat. Rev. Genet. *11*, 597–610.

Lehrbach, N.J., Armisen, J., Lightfoot, H.L., Murfitt, K.J., Bugaut, A., Balasubramanian, S., and Miska, E.A. (2009). LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in Caenorhabditis elegans. Nat. Struct. Mol. Biol. *16*, 1016–1020.

Liang, L., Wong, C.M., Ying, Q., Fan, D.N., Huang, S., Ding, J., Yao, J., Yan, M., Li, J., Yao, M., et al. (2010). MicroRNA-125b suppressesed human liver cancer cell proliferation and metastasis by directly targeting oncogene LIN28B2. Hepatology *52*, 1731–1740.

Liu, J., Carmell, M.A., Rivas, F.V., Marsden, C.G., Thomson, J.M., Song, J.J., Hammond, S.M., Joshua-Tor, L., and Hannon, G.J. (2004). Argonaute2 is the catalytic engine of mammalian RNAi. Science *305*, 1437–1441.

Lu, L., Katsaros, D., Shaverdashvili, K., Qian, B., Wu, Y., de la Longrais, I.A., Preti, M., Menato, G., and Yu, H. (2009). Pluripotent factor lin-28 and its homologue lin-28b in epithelial ovarian cancer and their associations with disease outcomes and expression of let-7a and IGF-II. Eur. J. Cancer *45*, 2212–2218.

Martinez, N.J., and Gregory, R.I. (2010). MicroRNA gene regulatory pathways in the establishment and maintenance of ESC identity. Cell Stem Cell *7*, 31–35.

Matera, A.G., Terns, R.M., and Terns, M.P. (2007). Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat. Rev. Mol. Cell Biol. *8*, 209–220.

Melton, C., Judson, R.L., and Blelloch, R. (2010). Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. Nature *463*, 621–626.

Moss, E.G., Lee, R.C., and Ambros, V. (1997). The cold shock domain protein LIN-28 controls developmental timing in C. elegans and is regulated by the lin-4 RNA. Cell *88*, 637–646.

Newman, M.A., Thomson, J.M., and Hammond, S.M. (2008). Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. RNA *14*, 1539–1549.

Oh, J.S., Kim, J.J., Byun, J.Y., and Kim, I.A. (2010). Lin28-let7 modulates radiosensitivity of human cancer cells with activation of K-Ras. Int. J. Radiat. Oncol. Biol. Physiol. *76*, 5–8.

Peng, S., Maihle, N.J., and Huang, Y. (2010). Pluripotency factors Lin28 and Oct4 identify a sub-population of stem cell-like cells in ovarian cancer. Oncogene *29*, 2153–2159.

Piskounova, E., Viswanathan, S.R., Janas, M., LaPierre, R.J., Daley, G.Q., Sliz, P., and Gregory, R.I. (2008). Determinants of microRNA processing inhibition by the developmentally regulated RNA-binding protein Lin28. J. Biol. Chem. *283*, 21310–21314.

Rybak, A., Fuchs, H., Smirnova, L., Brandt, C., Pohl, E.E., Nitsch, R., and Wulczyn, F.G. (2008). A feedback loop comprising lin-28 and let-7 controls pre-let-7 maturation during neural stem-cell commitment. Nat. Cell Biol. *10*, 987–993.

Ryder, S.P., Recht, M.I., and Williamson, J.R. (2008). Quantitative analysis of protein-RNA interactions by gel mobility shift. Methods Mol. Biol. *488*, 99–115.

Scott, M.S., Boisvert, F.M., McDowall, M.D., Lamond, A.I., and Barton, G.J. (2010). Characterization and prediction of protein nucleolar localization sequences. Nucleic Acids Res. *38*, 7388–7399.

Shell, S., Park, S.M., Radjabi, A.R., Schickel, R., Kistner, E.O., Jewell, D.A., Feig, C., Lengyel, E., and Peter, M.E. (2007). Let-7 expression defines two differentiation stages of cancer. Proc. Natl. Acad. Sci. USA *104*, 11400–11405.

Slack, F. (2009). let-7 microRNA reduces tumor growth. Cell Cycle *8*, 1823.

Small, E.M., and Olson, E.N. (2011). Pervasive roles of microRNAs in cardiovascular biology. Nature *469*, 336–342.

Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., Harano, T., Yatabe, Y., Nagino, M., Nimura, Y., et al. (2004). Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res. *64*, 3753–3756.

Trang, P., Medina, P.P., Wiggins, J.F., Ruffino, L., Kelnar, K., Omotola, M., Homer, R., Brown, D., Bader, A.G., Weidhaas, J.B., and Slack, F.J. (2010). Regression of murine lung tumors by the let-7 microRNA. Oncogene *29*, 1580–1587.

Van Wynsberghe, P.M., Kai, Z.S., Massirer, K.B., Burton, V.H., Yeo, G.W., and Pasquinelli, A.E. (2011). LIN-28 co-transcriptionally binds primary let-7 to regulate miRNA maturation in Caenorhabditis elegans. Nat. Struct. Mol. Biol. *18*, 302–308.

Viswanathan, S.R., and Daley, G.Q. (2010). Lin28: A microRNA regulator with a macro role. Cell *140*, 445–449.

Viswanathan, S.R., Daley, G.Q., and Gregory, R.I. (2008). Selective blockade of microRNA processing by Lin28. Science *320*, 97–100.

Viswanathan, S.R., Powers, J.T., Einhorn, W., Hoshida, Y., Ng, T.L., Toffanin, S., O'Sullivan, M., Lu, J., Phillips, L.A., Lockhart, V.L., et al. (2009). Lin28 promotes transformation and is associated with advanced human malignancies. Nat. Genet. *41*, 843–848.

Wang, Y.C., Chen, Y.L., Yuan, R.H., Pan, H.W., Yang, W.C., Hsu, H.C., and Jeng, Y.M. (2010). Lin-28B expression promotes transformation and invasion in human hepatocellular carcinoma. Carcinogenesis *31*, 1516–1522.

West, J.A., Viswanathan, S.R., Yabuuchi, A., Cunniff, K., Takeuchi, A., Park, I.H., Sero, J.E., Zhu, H., Perez-Atayde, A., Frazier, A.L., et al. (2009). A role for Lin28 in primordial germ-cell development and germ-cell malignancy. Nature *460*, 909–913.

Yang, X., Lin, X., Zhong, X., Kaur, S., Li, N., Liang, S., Lassus, H., Wang, L., Katsaros, D., Montone, K., et al. (2010). Double-negative feedback loop between reprogramming factor LIN28 and microRNA let-7 regulates aldehyde dehydrogenase 1-positive cancer stem cells. Cancer Res. *70*, 9463–9472.

Yu, F., Yao, H., Zhu, P., Zhang, X., Pan, Q., Gong, C., Huang, Y., Hu, X., Su, F., Lieberman, J., and Song, E. (2007a). let-7 regulates self renewal and tumorigenicity of breast cancer cells. Cell *131*, 1109–1123.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007b). Induced pluripotent stem cell lines derived from human somatic cells. Science *318*, 1917–1920.

Zhu, H., Shah, S., Shyh-Chang, N., Shinoda, G., Einhorn, W.S., Viswanathan, S.R., Takeuchi, A., Grasemann, C., Rinn, J.L., Lopez, M.F., et al. (2010). Lin28a transgenic mice manifest size and puberty phenotypes identified in human genetic association studies. Nat. Genet. *42*, 626–630.

LETTER

doi:10.1038/nature12119

A role for the Perlman syndrome exonuclease Dis3l2 in the Lin28–let-7 pathway

Hao-Ming Chang[1,2,3], Robinson Triboulet[1,2,3], James E. Thornton[1,2,3] & Richard I. Gregory[1,2,3]

The pluripotency factor Lin28 blocks the expression of let-7 microRNAs in undifferentiated cells during development, and functions as an oncogene in a subset of cancers[1]. Lin28 binds to let-7 precursor (pre-let-7) RNAs and recruits 3′ terminal uridylyl transferases to selectively inhibit let-7 biogenesis[2–4]. Uridylated pre-let-7 is refractory to processing by Dicer, and is rapidly degraded by an unknown RNase[5]. Here we identify Dis3l2 as the 3′–5′ exonuclease responsible for the decay of uridylated pre-let-7 in mouse embryonic stem cells. Biochemical reconstitution assays show that 3′ oligouridylation stimulates Dis3l2 activity *in vitro*, and knockdown of Dis3l2 in mouse embryonic stem cells leads to the stabilization of pre-let-7. Our study establishes 3′ oligouridylation as an RNA decay signal for Dis3l2, and identifies the first physiological RNA substrate of this new exonuclease, which is mutated in the Perlman syndrome of fetal overgrowth and causes a predisposition to Wilms' tumour development[6].

Post-transcriptional gene regulation by microRNAs (miRNAs) affects many developmental and physiological processes. Functioning by base-pairing with target messenger RNAs of complementary sequence, these ~22-nucleotide RNAs recruit the miRNA-induced silencing complex for translational repression and mRNA decay. Of particular relevance is the ancient let-7 family of miRNAs that is essential for normal development of *Caenorhabditis elegans*. Loss of the tumour suppressor function of the let-7 family affects various human cancers[7]. Expression of let-7 is dynamically regulated during development by the paralogous RNA-binding proteins Lin28A and Lin28B (refs 5, 8–10). Lin28 was identified as a regulator of developmental timing in worms, and has more recently been linked in mammals with controlling developmental timing and growth, as well as maintaining glucose homeostasis[1,11–13]. Lin28 is a pluripotency factor in stem cells, in which its expression helps to maintain an undifferentiated and proliferative state by blocking let-7 expression[14,15]. The Lin28–let-7 pathway is normally silenced in adult somatic cells, yet the expression of LIN28A or LIN28B is associated with a wide variety of human cancers[16,17]. Inhibition of this oncogenic pathway blocks the tumorigenicity of cancer cells[16].

Recent work has provided insight into the mechanisms underlying the Lin28-mediated selective regulation of let-7 (ref. 18). Lin28A functions in the cell cytoplasm, where it recruits the 3′ terminal uridylyl transferase (TUTase) Zcchc11 (also known as TUT4) or Zcchc6 (TUT7), which adds an oligouridine tail to pre-let-7 to inhibit Dicer processing and serves as a signal for the rapid decay of the uridylated RNA by an unknown nuclease[2–5,16]. We sought to identify the downstream nuclease(s), and used a biochemical approach to isolate factors that specifically associate with uridylated pre-let-7 (Fig. 1a, Supplementary Fig. 1 and Supplementary Table 1). This analysis showed that Lin28A and Zcchc11 were associated with both pre-let-7a-1 and uridylated pre-let-7a-1 (pre-let-7+14U), whereas Dis3l2, a 3′–5′ exonuclease, was specifically detected in the pre-let-7a-1+14U purification[6] (Fig. 1b). These mass spectrometry data were confirmed by western blot (Fig. 1c). Co-immunoprecipitation assays using a mouse embryonic stem-cell (ESC) line expressing a doxycycline (Dox)-inducible Flag–Lin28A transgene showed that Zcchc11, Zcchc6 and Dis3l2 are detectable by western blot in the Flag–Lin28A affinity eluate (Fig. 1d). Additional co-immunoprecipitation with Flag–TRBP, a cytoplasmic miRNA-binding protein, and Flag–Trim71, a cytoplasmic RNA-binding protein highly expressed in ESCs, confirmed the specificity of the Dis3l2 association with Lin28A (refs 15, 19) (Fig. 1e). To address whether this association is mediated through RNA, we performed further co-immunoprecipitation assays using either wild-type Lin28A or a mutant Lin28A (Lin28A(Trp46Ala)) that exhibits compromised RNA-binding activity towards pre-let-7 (ref. 18). We found less Dis3l2 associated with mutant Lin28A, and this association is strongly reduced after RNase A treatment (Fig. 1f). Overall, these results indicate that Lin28A associates with Dis3l2 in an RNA-dependent manner, and implicate Dis3l2 as a possible nuclease in the Lin28–let-7 pathway.

We next cloned and sequenced *Dis3l2* complementary DNA from mouse V6.5 ESCs. *Dis3l2.2* was confirmed as the major transcript variant expressed in V6.5 mouse ESCs that encodes an 870-amino-acid protein (Fig. 1b and Supplementary Fig. 2). Because pre-let-7 degradation occurs in the cell cytoplasm, we next examined the subcellular localization of Dis3l2 in ESCs (Supplementary Fig. 3). Dis3l2 was found to localize primarily to the cytoplasm of V6.5 ESCs, consistent with other cell types[6].

We carried out RNA degradation assays using affinity-purified Flag–Dis3l2. Dis3l2 was found to degrade pre-let-7+14U preferentially compared with pre-let-7 or an unrelated pre-miR-21 (Fig. 2a). To rule out the possibility that this observed activity was due to a co-purifying nuclease in the Flag–Dis3l2 immunoprecipitate, we generated a mutant Dis3l2-expressing construct by replacing a conserved residue in the catalytic domain (Asp389Asn) (Fig. 1b). Mutation of the equivalent aspartic acid in yeast Dis3(Asp551Asn) abolishes exonuclease activity without interfering with RNA binding[20]. Indeed, Dis3l2(Asp389Asn) displayed no activity, whereas wild-type Dis3l2 displayed preferential activity towards pre-let-7+14U (Fig. 2b, c).

Next, to determine whether Dis3l2 is sufficient for the selective degradation of uridylated pre-let-7, we generated recombinant Dis3l2 (rDis3l2) protein in *Escherichia coli* (Fig. 2d). Although the activity of rDis3l2 was lower than that of the affinity-purified Flag–Dis3l2 complexes, we observed a similar preference for pre-let-7+14U compared with pre-let-7 using the rDis3l2 protein (Fig. 2e, f). To rule out the possibility that this observed activity was due to a bacterial nuclease that might co-purify with histidine-tagged Dis3l2 (His–Dis3l2), we generated a mutant (Asp389Asn) rDis3l2 and confirmed that this catalytic mutant displayed no RNase activity in these assays (Supplementary Fig. 4a). Considering that uridylated pre-let-7 was previously found to be associated with isolated Flag–Lin28A complexes, together with our co-immunoprecipitation data, we next explored whether Lin28A protein influenced Dis3l2 activity *in vitro*[5]. This analysis showed that Lin28A had no effect on Dis3l2 activity (Supplementary Fig. 5). Next, to measure the substrate preference of Dis3l2 for uridylated pre-let-7 quantitatively, we performed time course experiments with rDis3l2. These showed a strong preference for the

[1]Stem Cell Program, Boston Children's Hospital, Massachusetts 02115, USA. [2]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA. [3]Harvard Stem Cell Institute, Boston, Massachusetts 02115, USA.

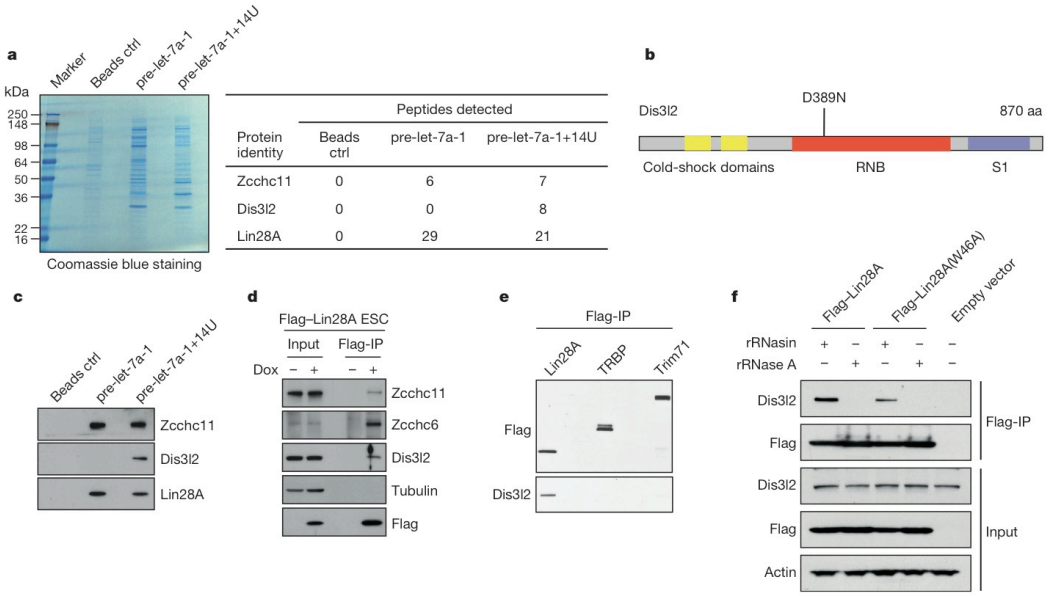


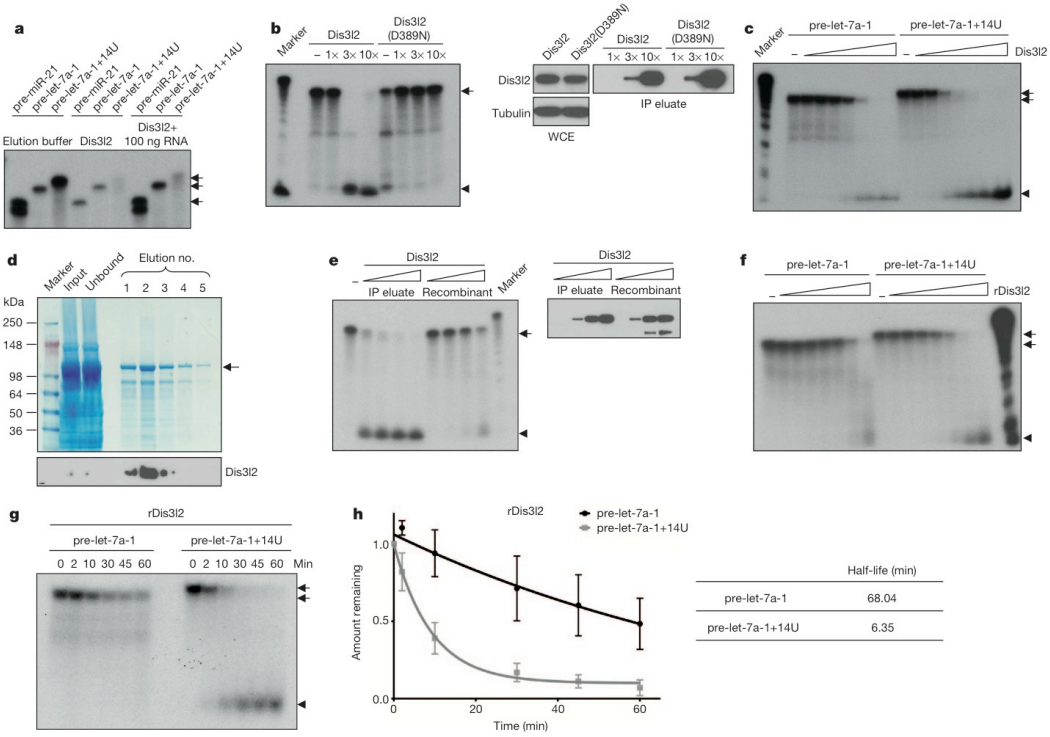
Figure 1 | Dis3l2 is associated with uridylated pre-let-7. a, Affinity-purified proteins analysed by Coomassie blue staining and mass spectrometry. Ctrl, control. **b**, Diagramatic representation of Dis3l2 protein (NCBI accession number NP_705758.1). Cold-shock domains, ribonuclease II domain (RNB) and S1 RNA-binding domain are indicated. The mutated catalytic Asp is indicated. aa, amino acids. **c**, Western blotting analysis of samples in **a** with indicated antibodies. **d**–**f**, Co-immunoprecipitation (IP) and western blot analyses. rRNasin, RNase inhibitor.

degradation of uridylated pre-let-7 compared to pre-let-7, with a more than tenfold difference in the relative RNA stability in these assays (Fig. 2g, h). Altogether, these results show that purified rDis3l2 preferentially degrades uridylated pre-let-7 *in vitro*, and that the oligouridine tail serves as a decay signal for this RNase.

To explore further the functional relationship between Lin28A, TUTase and Dis3l2-mediated RNA degradation, we next performed *in vitro* reconstitution assays. We showed previously that Lin28A enhances the uridylation activity of Zcchc11 towards pre-let-7. In these assays TUTase activity was measured by the incorporation of radiolabelled UTP[3,4,16]. However, owing to a limiting UTP concentration, the oligouridine tails added in these reactions are short, comprising only a few nucleotides. Supplementing such reactions with additional (non-radiolabelled) UTP leads to the generation of longer uridine tails (Fig. 3a, compare lanes 1, 2 and 3). Notably, the addition of Dis3l2 leads to the selective degradation of the pre-let-7 with longer uridine tails (Fig. 3a). These data define the minimal set of proteins and enzymes required to recapitulate the selective degradation of pre-let-7 observed *in vivo*, and raise questions about uridine tail-length requirements to stimulate Dis3l2-mediated degradation. To address this we prepared a panel of pre-let-7 RNA substrates with varying uridine tail lengths and monitored Dis3l2 degradation activity. Tails of at least 10 uridines were found to stimulate Dis3l2 activity, with maximal stimulation observed with uridine tails of 14 or greater (Fig. 3b). This result is consistent with the average length of the uridine tail found on pre-let-7 RNAs cloned and sequenced from Lin28-expressing cells[5].

To examine the Dis3l2 domain requirements, we generated three deletion mutants lacking the amino terminus, the carboxy terminus, or both the N- and C-terminal regions (Fig. 3c). RNA degradation assays and electromobility shift assays (EMSAs) revealed that truncation of either the N- or the C-terminal region abrogated both Dis3l2 binding and nuclease activities on uridylated pre-let-7 RNA (Fig. 3d, e). This suggests that both the cold-shock domains and the S1 domain are required for binding to uridylated pre-let-7. Although the catalytic mutant (Asp389Asn) Dis3l2 was inactive in RNA degradation assays, it retained the ability to bind to uridylated pre-let-7 selectively (Fig. 3e and Supplementary Fig. 4b). RNA degradation and binding assays using an unrelated RNA, pre-miR-21 with or without 14 uridines, established the sufficiency of an oligouridine tail to serve as a signal to trigger Dis3l2-mediated decay (Supplementary Fig. 4c, d).

To examine the role of Dis3l2 in the let-7 pathway, we used short interfering RNAs (siRNAs) to deplete Lin28A, Zcchc11 or Dis3l2 expression in mouse ESCs. We also included siRNAs that target the related family member Dis3l1 (Fig. 4a, b). We monitored the effects of gene knockdown on mature miRNA expression by quantitative reverse transcriptase PCR (qRT–PCR) and northern blot (Fig. 4c, d). Whereas Lin28A knockdown caused the expected accumulation of multiple let-7 miRNAs, knockdown of Dis3l2 or Dis3l1 had no effect. We did, however, observe a modest increase in let-7 expression in the Zcchc11-depleted samples, as previously reported[3–5]. Uridylated pre-let-7 has been previously shown to be resistant to cleavage by affinity-purified Dicer complexes, and consistently we found pre-let-7+14U to be a poor substrate for recombinant Dicer in processing assays. By comparison, pre-let-7 was processed by Dicer to ~22-nucleotide duplexes (Fig. 4e). Because pre-let-7+14U is inefficiently processed by Dicer, we postulated that knockdown of Dis3l2 in cells could lead to accumulation of uridylated pre-let-7 without affecting levels of mature let-7. To test this we developed a sensitive qRT–PCR-based assay for the specific detection of uridylated pre-let-7. We used an oligo(dA) primer for the reverse transcriptase first-strand cDNA synthesis step, and used primers complementary to pre-let-7 for the detection of uridylated pre-let-7 by quantitative PCR. This approach allowed us to detect uridylated pre-let-7 specifically (Supplementary Fig. 6a). RNA from the knockdown samples was size fractionated to measure relative levels of uridylated pre-let-7 in the <200-nucleotide fraction and the corresponding primary let-7 (pri-let-7) transcripts in the large



	Half-life (min)
pre-let-7a-1	68.04
pre-let-7a-1+14U	6.35

Figure 2 | Dis3l2 preferentially degrades uridylated pre-let-7. **a**, RNA degradation assay with Flag–Dis3l2 incubated with different radiolabelled pre-miRNAs. Where indicated, 100 ng competitor RNA was added to reduce background activity. **b**, RNA degradation assay with Flag-tagged Dis3l2 or mutant Dis3l2 (Dis3l2(D389N)) incubated with pre-let-7a-1+14U. WCE, whole-cell extract. **c**, RNA degradation assay with a titration of Flag–Dis3l2. **d**, His–Dis3l2 examined by Coomassie blue staining and western blot. **e**, Flag–Dis3l2 and His–Dis3l2 analysed by western blot and examined for activity. **f**, RNA degradation assay with a titration of His–Dis3l2. **g**, Representative time course assay. **h**, Quantification of three independent experiments as in **g** with the corresponding calculated RNA half-lives. $P < 0.01$ (two-way analysis of variance (ANOVA) test). The error bars represent the s.d. calculated from three independent experiments. In all relevant panels, arrows indicate radiolabelled pre-miRNA, arrowheads indicate degradation products.

>200-nucleotide fraction. This revealed the specific accumulation of uridylated pre-let-7a-1 and pre-let-7g after Dis3l2 knockdown (Fig. 4f), whereas levels of the corresponding pri-let-7 transcripts were unchanged (Fig. 4g). PCR products were cloned and sequenced to confirm the specificity of this assay (Supplementary Fig. 6b, c). Similar results were found using stable Dis3l2 knockdown ESCs (Supplementary Fig. 7a–e). To confirm further the role of Dis3l2 in the regulation of uridylated pre-let-7 levels, we performed northern blot assays using a probe complementary to the terminal loop region of let-7g. This revealed a slower-migrating pre-let-7 band after Dis3l2 depletion that probably corresponds to oligouridylated pre-let-7 (Fig. 4h, i). Finally, we individually depleted two additional 3′–5′ exonucleases, Exosc10 (also known as RRP6) and Rrp44 (Dis3), from ESCs and measured the relative uridylated pre-let-7 levels by qRT–PCR. Knockdown of these exosome-associated nucleases did not affect uridylated pre-let-7 levels (Supplementary Fig. 8). These results provide strong support that Dis3l2 is the downstream nuclease that mediates the decay of uridylated pre-let-7 in the Lin28 pathway.

Our results identify Dis3l2 as a new component of the Lin28–let-7 pathway, and as the downstream nuclease responsible for the decay of uridylated pre-let-7 (Supplementary Fig. 9). This contention is based on the following. First, Dis3l2 specifically associates with uridylated pre-let-7 in RNA affinity purifications and is detected as a component of a Lin28A-containing ribonucleoprotein complex(es). Second, purified Dis3l2 (but not catalytically inactive mutant Dis3l2) complexes display substrate preference for uridylated pre-let-7 in RNA degradation assays *in vitro*. Third, *in vitro* reconstitution experiments with recombinant Dis3l2 reveal the sufficiency of this enzyme for the preferential degradation of uridylated pre-let-7. Last, knockdown of Dis3l2 causes the accumulation of uridylated pre-let-7 in mouse ESCs.

Dis3l2 belongs to a family of related 3′–5′ exonucleases with similar domain organization to bacterial RNase II (refs 6, 21–23). Notably, germline mutations in the human *DIS3L2* gene were recently found to be responsible for Perlman syndrome, a rare, autosomal recessive, fetal overgrowth syndrome[6]. In addition to being large, affected individuals are hypotonic, have organomegaly, characteristic facial dysmorphism, renal abnormalities, neurodevelopmental problems, and a markedly high susceptibility to Wilms' tumours (nephroblastoma), with >60% of surviving children developing Wilms' tumours. Moreover, *DIS3L2* was found to be mutated in ~30% of sporadic Wilms' tumours analysed[6]. It will be important to explain the role of DIS3L2 in the genesis and development of Perlman syndrome and Wilms' tumours, and future experiments with knockout mouse models will shed light on this question. Our work uncovers the first, to our knowledge, physiological RNA substrate of Dis3l2. Considering the similarities between the disease phenotypes associated with Dis3l2 deletion and those caused

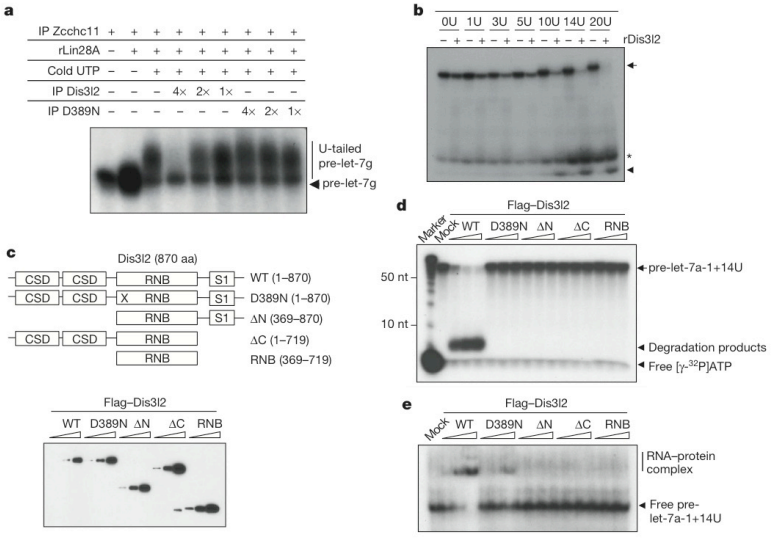


Figure 3 | Molecular determinants of Dis3l2 activity. **a**, Reconstitution assays reveal longer uridine tails are preferred substrates for Dis3l2. **b**, RNA degradation assays using *in vitro* transcribed pre-let-7 RNAs with indicated 3′ uridine (U) tail. Asterisk denotes nonspecific background RNA. **c**–**e**, Schematic representation and western blot of different Dis3l2 truncations (**c**) used for RNA degradation assays (**d**) and EMSAs (**e**). CSD, cold-shock domain; nt, nucleotide. Arrows and arrowheads are as in Fig. 2.

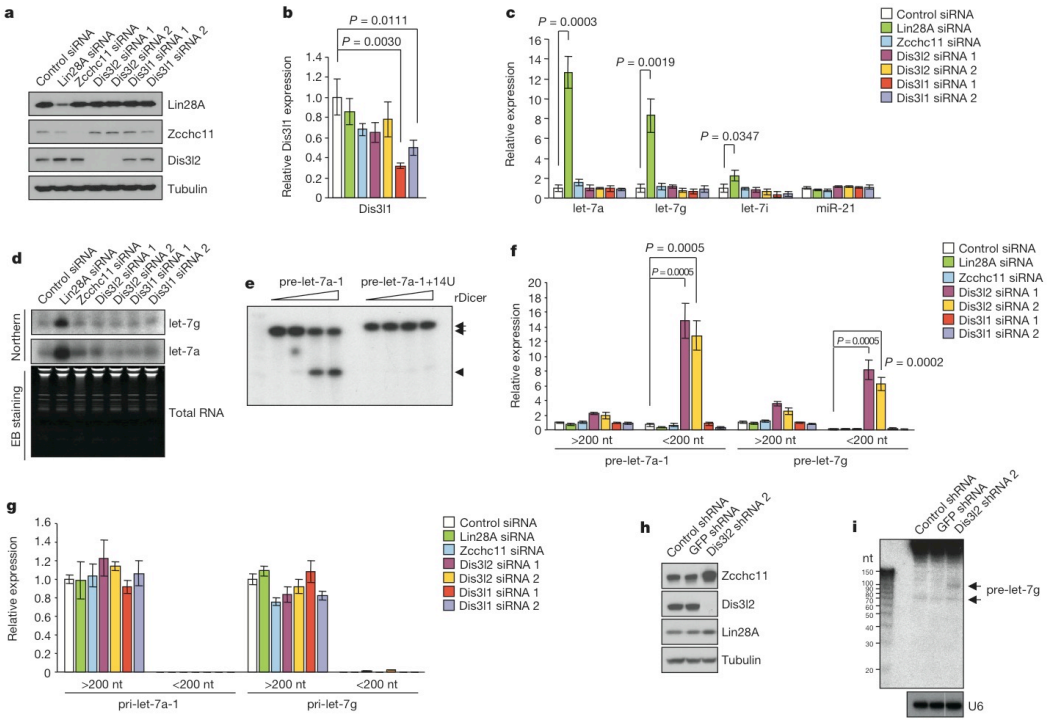


Figure 4 | Dis3l2 is required for degradation of uridylated pre-let-7 in ESCs. **a**, Western blot analysis of siRNA knockdowns. **b**, qRT–PCR analysis of Dis3l1 knockdown. Error bars denote s.d. ($n = 3$). **c**, Mature miRNA levels measured by qRT–PCR. Error bars denote s.d. ($n = 3$). **d**, Northern blot. EB, ethidium bromide. **e**, Dicer processing assay, with arrows indicating pre-miRNA and arrowhead indicating Dicer products. **f**, **g**, Total RNA from the samples in **a** was fractionated into >200-nucleotide- and <200-nucleotide-long RNA, and relative levels of uridylated pre-let-7 (**f**) or pri-let-7 (**g**) RNA were quantified by qRT–PCR. Error bars denote s.d. ($n = 3$). **h**, Western blot analysis of Dis3l2 cells. **i**, Northern blot analysis of pre-let-7g.

by Lin28 gain of function (that is, overgrowth and tumorigenesis), it is tempting to speculate that this new role of Dis3l2 in the Lin28–let-7 pathway is relevant to Perlman syndrome and cancer.

Our identification of a decay pathway for uridylated RNAs raises questions about how widespread this type of regulation might be on a transcriptome scale, as well as the mechanism by which oligouridylation promotes Dis3l2 ribonucleolytic activity. So far, there are few known examples in which 3′ uridylation can serve as a decay signal; these include histone mRNA regulation during the mammalian cell cycle, and the widespread uridylation-dependent mRNA decapping and decay in *Schizosaccharomyces pombe*[24–26]. Similarly, in *Saccharomyces cerevisiae*, the Trf4/Air2/Mtr4 polyadenylation (TRAMP) complex catalyses the addition of an oligoadenine tail that promotes 3′–5′ RNA decay by the exosome as part of a nuclear RNA surveillance mechanism[27,28]. In the case of pre-let-7, the 3′ oligouridylation has two consequences: to block Dicer processing, and to stimulate decay by Dis3l2. This two-step mechanism safeguards against the production of mature let-7 miRNA.

Note added in proof: Notably, the *S. pombe* Dis3l2 homologue displays preference for degradation of uridylated RNAs *in vitro*, and its deletion leads to accumulation of uridylated mRNAs, indicating that this regulatory pathway is more widespread (ref. 29).

METHODS SUMMARY

Affinity pull-down assays. Synthetic pre-let-7a-1 or pre-let-7a-1+14U was conjugated to agarose beads and incubated with whole-cell extract from P19 cells. Affinity eluate was subjected to SDS–PAGE and Coomassie blue staining. Bands were excised and subjected to mass spectrometric sequencing. Protein complexes were affinity-purified using anti-Flag M2 agarose beads (Sigma).

Plasmids and cDNA cloning. *Dis3l2* cDNA was cloned into pFlag-CMV2 (Sigma). Dis3l2(Asp389Asn) was generated by site-directed mutagenesis. cDNA was subcloned into pETDuet-1 for His-tagged Dis3l2 expression.

Recombinant Dis3l2 protein purification. Ni-NTA beads were used for the purification of His–Dis3l2 from IPTG (isopropyl-β-D-thiogalactoside)-induced BL21-CodonPlus competent bacteria (Stratagene).

RNA degradation assays. RNA degradation assays were performed using either 5′-end-labelled synthetic pre-miRNA or internally labelled *in vitro* transcribed pre-miRNA, together with Dis3l2 in RNA degradation buffer, and incubated at 37 °C.

RNA EMSAs. EMSA experiments were performed as described previously[16]. Nucleoprotein complexes were resolved by 4–20% non-denaturing TBE gel electrophoresis and visualized by autoradiography.

Transfections and siRNA/short hairpin RNA knockdowns. All transfections were performed with Lipofectamine 2000 (Invitrogen) as per manufacturer's instructions. Lentivirus production, infection and stable cell selection are as described[17].

qRT–PCR. RNA was isolated using TRIzol reagent (Invitrogen) and size fractionated using mirVana miRNA isolation kit (Ambion). For detection of uridylated pre-miRNA, size-fractionated RNA (<200-nucleotide RNA fraction) was treated with DNase then reverse transcribed using oligo$(dA)_{12}$ primer and SuperScript III (Invitrogen). qPCR was performed using iQ SYBR Green Supermix (Bio-Rad).

Northern blotting. Ten micrograms of total RNA from each sample was used for northern blotting, as previously described[19].

Dicer assays. Recombinant Flag–Dicer protein was purified from insect cells as previously described[19].

Full Methods and any associated references are available in the online version of the paper.

Received 17 October 2012; accepted 26 March 2013.
Published online 17 April 2013.

1. Thornton, J. E. & Gregory, R. I. How does Lin28 let-7 control development and disease? *Trends Cell Biol.* **22,** 474–482 (2012).
2. Heo, I. *et al.* TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell* **138,** 696–708 (2009).
3. Hagan, J. P., Piskounova, E. & Gregory, R. I. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nature Struct. Mol. Biol.* **16,** 1021–1025 (2009).
4. Thornton, J. E., Chang, H. M., Piskounova, E. & Gregory, R. I. Lin28-mediated control of let-7 microRNA expression by alternative TUTases Zcchc11 (TUT4) and Zcchc6 (TUT7). *RNA* **18,** 1875–1885 (2012).
5. Heo, I. *et al.* Lin28 mediates the terminal uridylation of let-7 precursor microRNA. *Mol. Cell* **32,** 276–284 (2008).
6. Astuti, D. *et al.* Germline mutations in *DIS3L2* cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. *Nature Genet.* **44,** 277–284 (2012).
7. Roush, S. & Slack, F. J. The let-7 family of microRNAs. *Trends Cell Biol.* **18,** 505–516 (2008).
8. Viswanathan, S. R., Daley, G. Q. & Gregory, R. I. Selective blockade of microRNA processing by Lin28. *Science* **320,** 97–100 (2008).
9. Newman, M. A., Thomson, J. M. & Hammond, S. M. Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA* **14,** 1539–1549 (2008).
10. Rybak, A. *et al.* A feedback loop comprising *lin-28* and *let-7* controls pre-*let-7* maturation during neural stem-cell commitment. *Nature Cell Biol.* **10,** 987–993 (2008).
11. Ambros, V. & Horvitz, H. R. Heterochronic mutants of the nematode *Caenorhabditis elegans*. *Science* **226,** 409–416 (1984).
12. Zhu, H. *et al. Lin28a* transgenic mice manifest size and puberty phenotypes identified in human genetic association studies. *Nature Genet.* **42,** 626–630 (2010).
13. Zhu, H. *et al.* The *Lin28/let-7* axis regulates glucose metabolism. *Cell* **147,** 81–94 (2011).
14. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318,** 1917–1920 (2007).
15. Chang, H. M. *et al.* Trim71 cooperates with microRNAs to repress Cdkn1a expression and promote embryonic stem cell proliferation. *Nature Commun.* **3,** 923 (2012).
16. Piskounova, E. *et al.* Lin28A and Lin28B inhibit let-7 microRNA biogenesis by distinct mechanisms. *Cell* **147,** 1066–1079 (2011).
17. Viswanathan, S. R. *et al.* Lin28 promotes transformation and is associated with advanced human malignancies. *Nature Genet.* **41,** 843–848 (2009).
18. Nam, Y., Chen, C., Gregory, R. I., Chou, J. J. & Sliz, P. Molecular basis for interaction of *let-7* microRNAs with Lin28. *Cell* **147,** 1080–1091 (2011).
19. Chendrimada, T. P. *et al.* TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436,** 740–744 (2005).
20. Dziembowski, A., Lorentzen, E., Conti, E. & Seraphin, B. A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nature Struct. Mol. Biol.* **14,** 15–22 (2007).
21. Tomecki, R. *et al.* The human core exosome interacts with differentially localized processive RNases: hDIS3 and hDIS3L. *EMBO J.* **29,** 2342–2357 (2010).
22. Staals, R. H. *et al.* Dis3-like 1: a novel exoribonuclease associated with the human exosome. *EMBO J.* **29,** 2358–2367 (2010).
23. Frazão, C. *et al.* Unravelling the dynamics of RNA degradation by ribonuclease II and its RNA-bound complex. *Nature* **443,** 110–114 (2006).
24. Mullen, T. E. & Marzluff, W. F. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5′ to 3′ and 3′ to 5′. *Genes Dev.* **22,** 50–65 (2008).
25. Norbury, C. J. 3′ Uridylation and the regulation of RNA function in the cytoplasm. *Biochem. Soc. Trans.* **38,** 1150–1153 (2010).
26. Rissland, O. S. & Norbury, C. J. Decapping is preceded by 3′ uridylation in a novel pathway of bulk mRNA turnover. *Nature Struct Mol. Biol.* **16,** 616–623 (2009).
27. LaCava, J. *et al.* RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* **121,** 713–724 (2005).
28. Wyers, F. *et al.* Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase. *Cell* **121,** 725–737 (2005).
29. Malecki, M. *et al.* The exoribonuclease Dis3L2 defines a novel eukaryotic RNA degradation pathway. *EMBO J.* http://dx.doi.org/10.1038/emboj.2013.63 (15 Mar, 2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank R. LaPierre for technical assistance and F.-L. Meng for Rrp44 and Exosc10 shRNA. Thanks to the Children's Hospital Boston/IDDRC Proteomics Core for mass spectrometry. R.I.G. was supported by grants from the US National Institute of General Medical Sciences (NIGMS) (R01GM086386) and The American Cancer Society (121635-RSG-11-175-01-RMC). J.E.T. was supported by a pre-doctoral fellowship from the National Science Foundation. R.T. was supported by the Wolbach fellowship from Boston Children's Hospital.

Author Contributions H.-M.C. designed and performed most of the experiments in Figs 1 and 2, and all of the experiments in Fig. 4. H.-M.C., R.T. and J.E.T. designed and performed experiments in Figs 1–3, as well as the Supplementary Figs. R.I.G. and H.-M.C. wrote the paper with input from R.T. and J.E.T.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.I.G. (rgregory@enders.tch.harvard.edu).

METHODS

Cell culture. HEK293 cells were maintained in DMEM, P19 cells in MEMα plus GlutaMax-1, and V6.5 and KH2 ESCs in DMEM with ESGRO (1,000 U ml^{-1}), supplemented with antibiotics, and 10% (for HEK293, P19) or 15% (for ESCs) FBS. A Dox-inducible Flag–Lin28A ESC line was used[3,12]. The MISSION short hairpin RNA (shRNA) plasmid DNA (Sigma; TRCN0000120760 for Dis3l2 shRNA 1, TRCN0000120761 for Dis3l2 shRNA 2, TRCN0000120745 for Dis3 shRNA, and TRCN0000123544 for Exosc10 shRNA) together with pLP1, pLP2 and VSVG were transfected into 293T cells to produce lentiviral particles that were used to infect V6.5 ESCs. The Dis3l2 shRNA stable cells were then created by puromycin (2.5 μg ml^{-1}) selection.

Affinity pull-down assays. For RNA affinity pull-down, synthetic mmu-pre-let-7a-1 or mmu-pre-let-7a-1+14U was conjugated to adipic acid dihydrazide agarose beads and incubated with whole-cell extract from P19 cells[8]. The affinity eluate was subjected to SDS–PAGE followed by Coomassie blue staining using the colloidal blue staining kit (Life Technologies). Bands were excised, and subjected to mass spectrometric sequencing. The sequencing results were further confirmed by western blotting. For co-immunoprecipitation assays using mouse KH2 ESCs expressing the Dox-inducible Flag–Lin28A transgene, cells were treated with Dox at 6 μg ml^{-1} for 48 h and then collected in lysis buffer (20 mM Tris-HCl, pH 8.0, 137 mM NaCl, 1 mM EDTA, 1% (v/v) Triton X-100, 10% (v/v) glycerol, 1.5 mM $MgCl_2$, 1 mM dithiothreitol (DTT) and 0.2 mM PMSF (phenylmethylsulphonyl fluoride)) supplemented with 40 U ml^{-1} of RNase inhibitor (rRNasin, Promega). Protein complexes were affinity-purified using anti-Flag M2 agarose beads (Sigma). Beads were washed with lysis buffer seven times before elution with 0.5 mg ml^{-1} Flag peptide. The eluates were analysed by SDS–PAGE and western blotting. For co-immunoprecipitation assays using ectopically expressed Flag–Lin28A, Flag–Lin28A(Trp46Ala), Flag–TRBP and Flag–Trim71, V6.5 ESCs were transfected using Lipofectamine 2000 (Invitrogen) and collected 48 h later. Cells were lysed as described above, except for the addition of RNase A (20 mg ml^{-1} final; QIAGEN) where indicated.

Mass spectrometry. The mass spectrometric protein analysis was performed at the Proteomics Center at Boston Children's Hospital. Bands of interest were excised from the Coomassie-stained SDS–PAGE gel, washed with a 2:1 ratio of 100 mM ammonium bicarbonate/acetonitrile, reduced with 10 mM DTT at 56 °C for 45 min, and alkylated for 30 min at room temperature, in the dark, with 55 mM iodoacetamide. Samples were digested with sequencing grade trypsin (Promega) at a concentration of 12.5 ng μl^{-1} in 100 mM ammonium bicarbonate at 37 °C overnight. Peptides were extracted with 100 mM ammonium bicarbonate and acetonitrile, and then dried in a speedvac. Samples were resuspended in 5% acetonitrile and 5% formic acid before direct injection into the liquid chromatography–mass spectrometry (LC–MS) system comprising of a nanoLC AS-2 autosampler, a nanoLC 2D HPLC pump (both Eksigent), and an LTQ mass spectrometer (Thermo Scientific). The liquid chromatography system for the mass spectrometer featured a reversed-phase column in-house packed into PicoTip Emitters (New Objective) using Magic C18 (3 μm, 200 Å; Michrom Bioresource) packing material. The peptides were eluted with a 30-min linear gradient, and data were acquired in a data-dependent fashion, that is, the six most abundant species were selected for fragmentation by collision induced dissociation. The .raw files were converted into .mgf files using in-house written scripts[30]. For each fragment ion spectrum, only the 200 most intense fragment ions were exported into the .mgf file. The mass spectrometric data were searched against the Uniprot Mouse database using the protein identification software Protein Pilot. The results were then filtered to include only proteins with a global false discovery rate of 1%.

Plasmids and DNA cloning. Mouse *Dis3l2* cDNA was generated by PCR using the forward (5′-AAGCTTGCGGCCGCGAACCATCCTGACTACAAGCTGAACCTTCGG-3′) and reverse (5′-AGACCTAGTCGACTCAGTCCTCAGGCTCCTCATCAGACGCC-3′) primers, and was cloned into the NotI and SalI sites of pFlag-CMV2 (Sigma). To generate the Dis3l2(Asp389Asn) mutant, site-directed mutagenesis was performed using the forward (5′-CTGCTCGCGACCTTAATGATGCCCTCGC-3′) and reverse (5′-GCGAGGGCATCATTAAGGTCGCGAGCAG-3′) primers. For generating His-tagged Dis3l2, the forward (5′-ACTAGGAATTCGAACCATCCTGACTACAAGCTGAACCTTCGG-3′) and reverse (5′-AAGCTTGCGGCCGCTCAGTCCTCAGGCTCCTCATCAGACGCC-3′) primers were used to amplifiy *Dis3l2* cDNA that was cloned into the EcoRI and NotI sites of pETDuet-1. For C- and N-terminal green fluorescent protein fusions (CT-GFP and NT-GFP, respectively), the GFP fusion TOPO TA expression kits (Invitrogen) were used. For CT-GFP fusions, the forward (5′-ACCATGAACCATCCTGACTACAAGCTGAAC-3′) and reverse (5′-CGTCCTCAGGCTCCTCATCAG-3′) primers were used. For NT-GFP fusions, the forward (5′-AACCATCCTGACTACAAGCTGAAC-3′) and reverse (5′-TCAGTCCTCAGGCTCCTCATCAG-3′) primers were used. For deletion mutants, Dis3l2-truncated cDNA was amplified by PCR, with the forward (5′-AACAAGCGGCCGCGAACCATCCTGACTACAAGCTGAACC-3′) and reverse (5′-AACAAGAATTGAGTAGCCCAGAGCAGCAGC-3′) primers to generate the C-terminal deletion mutant; with the forward (5′-AACAAGCGGCCGCGAGAAGAGACCTAAGGAAAGACTGTATCTTCAC-3′) and reverse (5′-AACAAGAATTCAGTCCTCAGGCTCCTCATC-3′) primers to generate the N-terminal deletion mutant; and with the forward (5′-AACAAGCGGCCGCGAGAAGAGACCTAAGGAAAGACTGTATCTTCAC-3′) and reverse (5′-AACAAGAATTGAGTAGCCCAGAGCAGCAGC-3′) primers to generate the N and C terminus deletion. These PCR products were cloned into NotI and EcoRI sites of the pFLAG-CMV2 vector (Sigma). Flag–TRBP and Flag–Trim71 constructs were as described[15,19].

Recombinant Dis3l2 protein purification. Transformed BL21-CodonPlus competent bacteria (Stratagene) were grown to an attenuance (*D*) at 600 nm of 0.4–0.6. Expression was induced with 100 μM IPTG (isopropyl-β-D-thiogalactoside) for 2–3 h. Cell pellets were resuspended in cold lysis buffer (20 mM imidazole in PBS, pH 8.0, and 0.1% PMSF) and sonicated. Cleared lysates were incubated with Ni-NTA beads and after 90 min incubation at 4 °C the beads were washed with 8 column volumes wash buffer (10 mM Tris, pH 7.8, 50 mM imidazole, pH 8.0, 500 mM NaCl and 0.1% PMSF). Bound His-tagged proteins were eluted from the column with 1 volume elution buffer (10 mM Tris, pH 7.8, 500 mM imidazole, pH 8.0, 500 mM NaCl and 0.1% fresh PMSF), and dialysed overnight against BC100 (20 mM Tris-HCl, pH 7.8, 100 mM KCl, 0.2 mM EDTA, 10% glycerol). Purified protein was dialysed against RNA degradation buffer (see below) and supplemented with 20% glycerol before storage at −80 °C.

RNA degradation assays. RNA degradation assays were performed in a total of 20 μl reaction using 6.25 nM 5′ end-labelled pre-miR-21, pre-let-7a-1 or pre-let-7a-1+14U RNA together with Dis3l2 and/or Lin28A. The reactions were set up in the RNA degradation buffer (20 mM HEPES-KOH, pH 7.5, 50 mM KCl, 0.05 mM $MgCl_2$ and 1 mM DTT), and incubated at 37 °C for 60 min. For time course assays, recombinant His–Dis3l2 was incubated with radiolabelled pre-let-7a-1 or pre-let-7a-1+14U. Bands from three independent experiments were quantified using ImageJ (NIH) and plotted using Prism (Graphpad). Values were fitted to one-phase decay curves with error bars representing s.d. ($n = 3$). *In vitro* uridylation assays were performed as described previously[3,4,16] with the addition of 10 μM cold uridine triphosphate and immunopurified Dis3l2 where indicated. The following synthetic RNA sequences were used (all from Dharmacon): mmu-pre-let-7a-1, 5′-UGAGGUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAACUCUACUGUCUUUCC-3′; mmu-pre-let-7a-1+14U, 5′-UGAGGUAGUAGGUUGUAUAGUUUUAGGGUCACACCCACCACUGGGAGAUAACUAUACAAUCUACUGUCUUUCCUUUUUUUUUUUUUU-3′; mmu-pre-miR-21, 5′-UAGCUUAUCAGACUGAUGUUGACUGUUGAAUCUCAUGGCAACAGCAGUCGAUGGGCUGUC-3′.

***In vitro* transcription of pre-miRNAs.** *In vitro* transcribed pre-let-7g RNAs were generated as substrates for RNA degradation assays (in Fig. 3b). DNA templates for *in vitro* transcription of pre-let-7 with different 3′ ends by PCR amplification were generated using a universal 5′ primer (ACGGTTCAGCTAATACGACTCACTATAGGGTGAGGTAGTAGTTTGTACAGTTTGAGG) (T7 promoter sequence underlined) and a 3′ primer (listed below) to amplify from a plasmid DNA template containing pri-let-7g (ref. 8). PCR products were cloned and sequence verified. DNA templates (PCR products) were gel-purified and *in vitro* transcription was performed according to Riboprobe *In-Vitro* Transcription Systems using [α-^{32}P]rGTP and T7 RNA polymerase (Promega). The labelled pre-miRNAs were treated with RQ1 DNase and cleaned by illustra MicroSpin G-25 Column (GE Healthcare Life Sciences).

The following 3′ primers for pre-let-7g-$(U)_n$ were used: pre-let-7g-0U, GCAAGGCAGTGGCCTGTACAGTTATC; pre-let-7g-1U, AGCAAGGCAGTGGCCTGTACAGTTATC; pre-let-7g-3U, AAAGCAAGGCAGTGGCCTGTACAGTTATC; pre-let-7g-5U, AAAAAGCAAGGCAGTGGCCTGTACAGTTATC; pre-let-7g-10U, AAAAAAAAAAGCAAGGCAGTGGCCTGTACAGTTATC; pre-let-7g-10U, AAAAAAAAAAGCAAGGCAGTGGCCTGTACAGTTATC; pre-let-7g-14U, AAAAAAAAAAAAAAGCAAGGCAGTGGCCTGTACAGTTATC; pre-let-7g-20U, AAAAAAAAAAAAAAAAAAAAGCAAGGCAGTGGCCTGTACAGTTATC; pre-miR-21 forward, ACGGTTCAGCTAATACGACTCACTATAGGGTAGCTTATCAGACTGATGTTGACTG; pre-miR-21 reverse, GACAGCCCATCGACTGCTGTTG; pre-miR-21-14U reverse, AAAAAAAAAAAAAAGACAGCCCATCGACTGCTGT.

RNA EMSAs. EMSA experiments were performed as described previously[16]. In brief, 1 nM of the indicated radiolabelled synthetic RNA was incubated in the binding buffer (50 mM Tris, pH 7.6, 100 mM NaCl, 10 mM β-mercaptoethanol and 1 U μl^{-1} rRNasin (Promega) with varying concentrations of catalytically inert recombinant His–Dis3l2 or recombinant His–Lin28 in the absence of competitor RNA. Nucleoprotein complexes were resolved by 4–20% non-denaturing TBE gel electrophoresis (Bio-Rad) and visualized by autoradiography.

Antibodies and synthetic RNA. The following antibodies and the working concentrations were used: anti-Dis3l2 (0.4 μg ml^{-1}; Novus Biologicals, NBP1-84740), anti-Zcchc11 (1:1,000; ProteinTech Group, 18980-1-AP), anti-Zcchc6 (1:500; Open Biosystems, custom-made), anti-Rrp44 (1:1,000; ProteinTech Group, 14689-1-AP), anti-Exosc10 (1 μg ml^{-1}; Abcam, ab50558), anti-Lin28A (A177) (1:1,000; Cell Signaling, 3978), anti-Flag (1:5,000; Sigma, A8592), and anti-tubulin (1:10,000; Abcam, ab6046).

Transfections and siRNA/shRNA knockdowns. All transfections were performed with Lipofectamine 2000 (Invitrogen) per manufacturer's instructions. The sequences of the shRNAs and siRNAs are listed below. Lentivirus production, infection and stable cell selection are as described[17]. The following synthetic RNA sequences were used (all from Dharmacon): control siRNA 1, 5′-UGGUUUACAUGUCGACUAA-3′; control siRNA 2, 5′-UGGUUUACAUGUUGUGUGA-3′; Lin28A siRNA, 5′-GGGUUGUGAUGACAGGCAA-3′; Zcchc11 siRNA, 5′-GGGCUAAGCUGUGCUAUAU-3′; mouse Dis3l2 siRNA 1, 5′-CCGCUUUGCUGACGUCAUA-3′; mouse Dis3l2 siRNA 2, 5′-GAAUUUACGUACCUCUCAA-3′; mouse Dis3l1 siRNA 1, 5′-AGGAACUACUGGACGGAAA-3′; mouse Dis3l1 siRNA 2, 5′-UGAAACAGAAGGCGUAUUU-3′.

mRNA and miRNA qRT–PCR. Total RNA was isolated using TRIzol reagent (Invitrogen). For fractionation, total RNA was processed by mirVana miRNA isolation kit according to the manufacturer's instructions (Ambion). For mRNA, 100 ng of total RNA was reverse transcribed using random hexamers and SuperScript III (Invitrogen). For mature miRNA, 10 ng of total RNA was reverse transcribed using gene-specific stem–loop reverse transcriptase primers and MultiScribe Reverse Transcriptase (Applied Biosystems). For pre-miRNA, <200-nucleotide fractionated RNA was first treated with 0.66 U RNase-free DNase (Promega) (60 min at 37 °C), stopped with 1 mM of EDTA (10 min at 65 °C), and reverse transcribed by oligo(dA)$_{12}$ (60 min at 50 °C) using SuperScript III (Invitrogen). The resulting cDNA was further digested with RNase H (30 min at 37 °C). For mRNA and pre-miRNA, iQ SYBR Green Supermix (Bio-Rad) was used for quantifying the cDNA. For mature miRNA, TaqMan universal PCR master mix, no AmpErase UNG (Applied Biosystems) was used for cDNA detection. All qPCR reactions were performed using iCycler iQ multicolor real-time PCR detection system (Bio-Rad). Normalization controls include mouse *Actb* for mRNAs as well as for pri-miRNAs, *U6* for pre-miRNA, and snoRNA142 for mature miRNAs. For all RT–PCRs, minus reverse transcriptase and water control samples were included, and in all cases the signals were undetectable (data not shown). The following primer sequences were used for qPCR: pre-let-7a-1, forward, 5′-TGAGGTAGTAGGTTGTATAGTTTTAGGG-3′; reverse, 5′-GGAAAGACAGTAGATTGTATAGTTATC-3′; pri-let-7a-1, forward, 5′-CTTTCAACATTCACCCTGGATGTTC-3′; reverse, 5′-GAGACCCCATGAATGCAGACTTT-3′; pre-let-7g, forward, 5′-TGAGGTAGTAGTTTGTACAGTTTGAGG-3′; reverse, 5′-GCAAGGCAGTGGCCTGTACAGTTATC-3′; pri-let-7g, forward, 5′-GTTCTCTTTTGCCTGATTCCAGG-3′; reverse, 5′-CATTTGGTAGCTGGTGCACTG-3′; *U6*, forward, 5′-CTCGCTTCGGCAGCACA-3′; reverse, 5′-AACGCTTCACGAATTTGCGT-3′; *Actb*, forward, 5′-CAGAAGGAGATTACTGCTCTGGCT-3′, reverse, 5′- TACTCCTGCTTGCTGATCCACATC-3′; *Dis3l1*, forward, 5′- AGTTGACAGACATAGCTCGCCACA-3′; reverse, 5′-TGGTTGGCTAGGATCATGCACTCA-3′.

Northern blotting. Ten micrograms of total RNA from each sample was used for northern blotting as previously described[19]. Probe sequences for detecting precursor and mature miRNA are as follows: pre-let-7g, 5′-TATCTCCTGTACCGGGTGGTATCATAGACCCTCA-3′; let-7a, 5′-AACTATACAACCTACTACCTCA-3′; let-7g, 5′-AACTGTACAAACTACTACCTCA-3′.

Dicer assays. Recombinant Flag–Dicer protein was purified from insect cells as previously described[19]. Dicer processing of pre-let-7 or pre-let-7+14U was performed by incubating recombinant Dicer with gel-purified 5′ end-labelled synthetic pre-miRNA in a buffer containing 3.2 mM $MgCl_2$, 20 mM Tris-HCl, pH 7.9, 0.1 M KCl, 10% glycerol, 5 mM DTT, 0.2 mM PMSF and 40 U ml^{-1} RNase inhibitor (RNasin, Promega) for 1 h at 37 °C. Samples were resolved by 15% denaturing polyacrylamide gel.

30. Renard, B. Y. *et al.* When less can yield more — computational preprocessing of MS/MS spectra for peptide identification. *Proteomics* **9,** 4978–4984 (2009).